

DMedium_10_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 | 17.12 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 2.1 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 33.2 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 65.3 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 87.0 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 53.2 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 23.7 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 7.20 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 20.1 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
|                                     | 195 kB 76.4 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 6.9 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 6.1 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 49.9 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 41.7 MB/s
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 70.5 MB/s
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Collecting plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 53.2 MB/s
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 49.5 MB/s
Requirement already satisfied: scikit-learn in
/usr/local/lib/python3.7/dist-packages (from sklearn==0.0->-r requirements.txt
(line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    | 880 kB 60.3 MB/s

```



```
size=895260
sha256=6c4b20a4f316b2d4b10fecad4433fc168b69500d885ec96fc0047bb79843e782
  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 | 18.38 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-fkfn17sd
Created temporary directory: /tmp/pip-req-tracker-aevalqq9
Initialized build tracking at /tmp/pip-req-tracker-aevalqq9
Created build tracker: /tmp/pip-req-tracker-aevalqq9
Entered build tracker: /tmp/pip-req-tracker-aevalqq9
```

```

Created temporary directory: /tmp/pip-install-w7dul8u5
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-c4xv4fdc
  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-
aevalqq9'
  Running setup.py (path:/tmp/pip-req-build-c4xv4fdc/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-eu121cc7
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-eu121cc7/apex.egg-info
  writing /tmp/pip-pip-egg-info-eu121cc7/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-eu121cc7/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-eu121cc7/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-eu121cc7/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-eu121cc7/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-c4xv4fdc 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-aevalqq9'
Created temporary directory: /tmp/pip-unpack-mpzoxk__
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-k9rxnqez
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-k9rxnqez
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''"/tmp/pip-req-build-c4xv4fdc/setup.py''';

```



```
__file__='''/tmp/pip-req-build-c4xv4fdc/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-k9rxnqez
```

```
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_pipelining_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_pipelining_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-k9rxnqez/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=311a6180ffa6532ec925f6c4f51fbf3b5e6d09f90639cf42b89b0a7548878bae
Stored in directory: /tmp/pip-ephem-wheel-cache-
fkfni7sd/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-aevalqq9'
/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 39.5 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                                     | 3.9 MB 49.7 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                                     | 79 kB 9.4 MB/s
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
    |                                     | 9.0 MB 52.3 MB/s
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 56.2 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting jmespath<2.0.0,>=0.7.1

```

```

Downloading jmespath-1.0.1-py3-none-any.whl (20 kB)
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)
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 17 kB/s
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 1014, done.
remote: Counting objects: 100% (84/84), done.
remote: Compressing objects: 100% (74/74), done.
remote: Total 1014 (delta 27), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (1014/1014), 254.12 MiB | 17.24 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/DMedium_10_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, 567kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 38.6MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 32.6MB/s]
Downloading: 100% 501M/501M [00:07<00:00, 67.5MB/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: 0.41527894139289856
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=0.2772277227722772, 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: 0.48303526639938354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.33766233766233766, f1=0.30769230769230765,
best_f1=0.30769230769230765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4279116094112396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.3010752688172043, f1=0.3010752688172043,
best_f1=0.30769230769230765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26288220286369324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.358974358974359, f1=0.3132530120481928,
best_f1=0.3132530120481928
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35750138759613037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.6666666666666666, f1=0.4878048780487805,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30397137999534607
/usr/local/lib/python3.7/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=0.3389830508474576, f1=0.3333333333333337,
best_f1=0.4878048780487805
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5238106846809387
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=0.7692307692307692, f1=0.7407407407407408,
best_f1=0.7407407407407408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4667981266975403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.782608695652174, f1=0.7272727272727273,
best_f1=0.7272727272727273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16508431732654572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8571428571428571, f1=0.7999999999999999,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10997205972671509
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8387096774193549, f1=0.7878787878787878,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03317750245332718
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.7586206896551724,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02472708187997341
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.8148148148148148,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01257377490401268
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.8148148148148148,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009183956310153008
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.8148148148148148,
best_f1=0.7999999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017433784902095795
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.8148148148148148,

```

```

best_f1=0.7999999999999999
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 94522.45it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.88
real_f1 = 0.8
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, 218.68it/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: 0.5671741366386414
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4705401659011841
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.459580659866333
step: 30, loss: 0.0988544151186943
step: 40, loss: 0.12931038439273834
step: 50, loss: 0.13948439061641693
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.1837298721075058
step: 70, loss: 0.013270504772663116
step: 80, loss: 0.06280394643545151
step: 90, loss: 0.16272136569023132
step: 100, loss: 0.09150248765945435
step: 110, loss: 0.03839397802948952
step: 120, loss: 0.05033928155899048
step: 130, loss: 0.021304737776517868
step: 140, loss: 0.009141020476818085
step: 150, loss: 0.23017199337482452
step: 160, loss: 0.005092805251479149
step: 170, loss: 0.006846219766885042
step: 180, loss: 0.005294640548527241
step: 190, loss: 0.045495178550481796
step: 200, loss: 0.1328253448009491
step: 210, loss: 0.045401431620121
step: 220, loss: 0.014205848798155785
step: 230, loss: 0.0017515893559902906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775280898876404, f1=0.9853107344632768,
best_f1=0.9853107344632768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016711128409951925
step: 10, loss: 0.035359691828489304
step: 20, loss: 0.0405314676463604
step: 30, loss: 0.006151751149445772

```

```

step: 40, loss: 0.017523206770420074
step: 50, loss: 0.0024460398126393557
step: 60, loss: 0.005685105454176664
step: 70, loss: 0.009414843283593655
step: 80, loss: 0.001325415913015604
step: 90, loss: 0.007795072626322508
step: 100, loss: 0.03840670734643936
step: 110, loss: 0.006043126340955496
step: 120, loss: 0.007645296864211559
step: 130, loss: 0.010894086211919785
step: 140, loss: 0.0036782289389520884
step: 150, loss: 0.17272979021072388
step: 160, loss: 0.009879056364297867
step: 170, loss: 0.0020719747990369797
step: 180, loss: 0.016616089269518852
step: 190, loss: 0.01913040317595005
step: 200, loss: 0.009067382663488388
step: 210, loss: 0.004027864895761013
step: 220, loss: 0.0017853259341791272
step: 230, loss: 0.0035493671894073486
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.9787234042553192, f1=0.9774774774774775,
best_f1=0.9774774774774775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007548601366579533
step: 10, loss: 0.0059134019538760185
step: 20, loss: 0.010488742031157017
step: 30, loss: 0.001118474523536861
step: 40, loss: 0.0018045245669782162
step: 50, loss: 0.03523661941289902
step: 60, loss: 0.010884348303079605
step: 70, loss: 0.0006559767061844468
step: 80, loss: 0.00170845293905586
step: 90, loss: 0.002956889569759369
step: 100, loss: 0.00699732406064868
step: 110, loss: 0.0015466890763491392
step: 120, loss: 0.012478020042181015
step: 130, loss: 0.13050460815429688
step: 140, loss: 0.001484916778281331
step: 150, loss: 0.08920168876647949

```

```

step: 160, loss: 0.005263463128358126
step: 170, loss: 0.0007299064309336245
step: 180, loss: 0.0027395060751587152
step: 190, loss: 0.014020569622516632
step: 200, loss: 0.014007572084665298
step: 210, loss: 0.0020844147074967623
step: 220, loss: 0.002259241882711649
step: 230, loss: 0.006540542468428612
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9796380090497738, f1=0.9807037457434733,
best_f1=0.9807037457434733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025158405769616365
step: 10, loss: 0.0019442837219685316
step: 20, loss: 0.0013930562417954206
step: 30, loss: 0.0012895468389615417
step: 40, loss: 0.022999174892902374
step: 50, loss: 0.004403779748827219
step: 60, loss: 0.0022647660225629807
step: 70, loss: 0.00798546988517046
step: 80, loss: 0.0019118236377835274
step: 90, loss: 0.07278670370578766
step: 100, loss: 0.0018754720222204924
step: 110, loss: 0.0005873796180821955
step: 120, loss: 0.0013971984153613448
step: 130, loss: 0.021659625694155693
step: 140, loss: 0.03444381058216095
step: 150, loss: 0.0004254930536262691
step: 160, loss: 0.0014052096521481872
step: 170, loss: 0.01781908981502056
step: 180, loss: 0.20357847213745117
step: 190, loss: 0.0016299267299473286
step: 200, loss: 0.0049929809756577015
step: 210, loss: 0.0009949503000825644
step: 220, loss: 0.0015590616967529058
step: 230, loss: 0.0007104436517693102
/usr/local/lib/python3.7/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=0.9753914988814317, f1=0.9750566893424036,

```

```

best_f1=0.9807037457434733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006465572398155928
step: 10, loss: 0.0028282543644309044
step: 20, loss: 0.09489333629608154
step: 30, loss: 0.0029504341073334217
step: 40, loss: 0.004939338658004999
step: 50, loss: 0.002213103463873267
step: 60, loss: 0.005140322260558605
step: 70, loss: 0.002388187451288104
step: 80, loss: 0.045779451727867126
step: 90, loss: 0.0364808551967144
step: 100, loss: 0.0009198882034979761
step: 110, loss: 0.0006893239333294332
step: 120, loss: 0.000854937476105988
step: 130, loss: 0.014102301560342312
step: 140, loss: 0.000556422455701977
step: 150, loss: 0.17862962186336517
step: 160, loss: 0.004082889296114445
step: 170, loss: 0.002995668910443783
step: 180, loss: 0.0007535041077062488
step: 190, loss: 0.0033306588884443045
step: 200, loss: 0.001058898400515318
step: 210, loss: 0.004999442026019096
step: 220, loss: 0.0014996587997302413
step: 230, loss: 0.0012577975867316127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9831649831649831,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004934347234666348
step: 10, loss: 0.004523595329374075
step: 20, loss: 0.001536040217615664
step: 30, loss: 0.0010911163408309221
step: 40, loss: 0.0008156581316143274
step: 50, loss: 0.0005181797896511853
step: 60, loss: 0.001688934862613678
step: 70, loss: 0.000752091349568218

```

```

step: 80, loss: 0.0020014073234051466
step: 90, loss: 0.004481547977775335
step: 100, loss: 0.00029847698169760406
step: 110, loss: 0.011614806950092316
step: 120, loss: 0.0007308440399356186
step: 130, loss: 0.0004168792220298201
step: 140, loss: 0.0003011714434251189
step: 150, loss: 0.00035058666253462434
step: 160, loss: 0.0008307126699946821
step: 170, loss: 0.00019790080841630697
step: 180, loss: 0.0001372411206830293
step: 190, loss: 0.0001161534819402732
step: 200, loss: 0.0005503600696101785
step: 210, loss: 0.00105635158251971
step: 220, loss: 0.002629162510856986
step: 230, loss: 0.0010385707719251513
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, f1=0.978675645342312,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014618081040680408
step: 10, loss: 0.000760428374633193
step: 20, loss: 0.0010304506868124008
step: 30, loss: 0.0010210563195869327
step: 40, loss: 0.0014152000658214092
step: 50, loss: 0.0037871296517550945
step: 60, loss: 0.0033857994712889194
step: 70, loss: 0.00029587832977995276
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 80, loss: 0.0008821403025649488
step: 90, loss: 0.0003374551597516984
step: 100, loss: 0.0003569337713997811
step: 110, loss: 0.00048260812764056027
step: 120, loss: 0.0008757172618061304
step: 130, loss: 0.002065324457362294
step: 140, loss: 0.000425264413934201
step: 150, loss: 0.0032569507602602243
step: 160, loss: 0.0002680346369743347
step: 170, loss: 0.00027429056353867054
step: 180, loss: 0.00037244451232254505

```



```

step: 190, loss: 0.0008072804193943739
step: 200, loss: 0.006117960438132286
step: 210, loss: 0.005289046559482813
step: 220, loss: 0.0012264109682291746
step: 230, loss: 0.002660752274096012
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9819819819819819, f1=0.9784335981838819,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032750703394412994
step: 10, loss: 0.00914386659860611
step: 20, loss: 0.0008727599633857608
step: 30, loss: 0.000891387346200645
step: 40, loss: 0.0005820276564918458
step: 50, loss: 0.0007989630103111267
step: 60, loss: 0.000533653364982456
step: 70, loss: 0.0003948990488424897
step: 80, loss: 0.007503471802920103
step: 90, loss: 0.0005235726130194962
step: 100, loss: 0.0007086523110046983
step: 110, loss: 0.030301976948976517
step: 120, loss: 0.0006340177496895194
step: 130, loss: 0.00334475701674819
step: 140, loss: 0.002094842493534088
step: 150, loss: 0.06673581153154373
step: 160, loss: 0.002589188516139984
step: 170, loss: 0.0016222017584368587
step: 180, loss: 0.0005265786312520504
step: 190, loss: 0.0004117939970456064
step: 200, loss: 0.000160046765813604
step: 210, loss: 0.000663054350297898
step: 220, loss: 0.00017437931091990322
step: 230, loss: 0.0001691470097284764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9865470852017937, f1=0.9797297297297298,
best_f1=0.9797297297297298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002082361897919327
step: 10, loss: 0.0004482937220018357
step: 20, loss: 0.0005713360733352602
step: 30, loss: 0.001621659961529076
step: 40, loss: 0.00033484090818092227
step: 50, loss: 0.0010890539269894361
step: 60, loss: 0.0002977137628477067
step: 70, loss: 0.04565931484103203
step: 80, loss: 0.0003197356127202511
step: 90, loss: 0.03976648300886154
step: 100, loss: 0.00013957181363366544
step: 110, loss: 0.0001441080094082281
step: 120, loss: 0.01801474764943123
step: 130, loss: 0.00047375046415254474
step: 140, loss: 0.0003978186286985874
step: 150, loss: 0.0004418205644469708
step: 160, loss: 0.00026112276827916503
step: 170, loss: 6.991311238380149e-05
step: 180, loss: 0.0005484263529069722
step: 190, loss: 7.057737093418837e-05
step: 200, loss: 0.00012473054812289774
step: 210, loss: 0.0014579021371901035
step: 220, loss: 6.214883615029976e-05
step: 230, loss: 4.01616380258929e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9876543209876544, f1=0.9819819819819819,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.765596142737195e-05
step: 10, loss: 0.0001212259303429164
step: 20, loss: 0.013183514587581158
step: 30, loss: 4.071449438924901e-05
step: 40, loss: 0.00037998182233422995
step: 50, loss: 7.351020030910149e-05
step: 60, loss: 4.345514389569871e-05
step: 70, loss: 0.0001375453284708783
step: 80, loss: 3.0144143238430843e-05
step: 90, loss: 4.328793511376716e-05
step: 100, loss: 4.57533351436723e-05

```

```

step: 110, loss: 9.893294918583706e-05
step: 120, loss: 4.3620497308438644e-05
step: 130, loss: 6.371299241436645e-05
step: 140, loss: 6.278883665800095e-05
step: 150, loss: 0.0011681392788887024
step: 160, loss: 3.0407927624764852e-05
step: 170, loss: 0.00018427932809572667
step: 180, loss: 0.00025361881125718355
step: 190, loss: 7.810199895175174e-05
step: 200, loss: 0.0002731051354203373
step: 210, loss: 0.0002499771071597934
step: 220, loss: 9.948650404112414e-05
step: 230, loss: 5.5616386816836894e-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=0.9865470852017937, f1=0.9819819819819819,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.805329758208245e-05
step: 10, loss: 4.9372200010111555e-05
step: 20, loss: 4.435507071320899e-05
step: 30, loss: 9.946752106770873e-05
step: 40, loss: 2.59116441156948e-05
step: 50, loss: 0.00011924448335776106
step: 60, loss: 0.009809069335460663
step: 70, loss: 5.0016151362797245e-05
step: 80, loss: 0.00029530812753364444
step: 90, loss: 0.07254774123430252
step: 100, loss: 0.0018962898757308722
step: 110, loss: 0.0021425362210720778
step: 120, loss: 0.0011209867661818862
step: 130, loss: 0.00013067368126939982
step: 140, loss: 0.001972750062122941
step: 150, loss: 0.00017516168009024113
step: 160, loss: 0.05906057730317116
step: 170, loss: 0.1909409761428833
step: 180, loss: 0.0036212848499417305
step: 190, loss: 0.0009418915142305195
step: 200, loss: 0.0038117466028779745
step: 210, loss: 0.0002288753166794777
step: 220, loss: 0.00030681482166983187
step: 230, loss: 0.0002854266786016524
/usr/local/lib/python3.7/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=0.9841628959276018,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023713089467491955
step: 10, loss: 0.0001262676960323006
step: 20, loss: 0.017535319551825523
step: 30, loss: 0.06477443873882294
step: 40, loss: 0.00013433618005365133
step: 50, loss: 0.0010076753096655011
step: 60, loss: 0.0009837363613769412
step: 70, loss: 0.00015762768452987075
step: 80, loss: 5.257554221316241e-05
step: 90, loss: 0.0007075613830238581
step: 100, loss: 0.00010825973004102707
step: 110, loss: 0.00012080513988621533
step: 120, loss: 0.00010827821824932471
step: 130, loss: 7.82811184762977e-05
step: 140, loss: 0.0001469717244617641
step: 150, loss: 0.00022138826898299158
step: 160, loss: 0.0018820398254320025
step: 170, loss: 0.00015448741032741964
step: 180, loss: 8.928826719056815e-05
step: 190, loss: 0.004429514054208994
step: 200, loss: 0.0003641980292741209
step: 210, loss: 0.00015619859914295375
step: 220, loss: 0.014160501770675182
step: 230, loss: 0.0014076537918299437
/usr/local/lib/python3.7/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=0.9853768278965129, f1=0.9831271091113611,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000281314249150455
step: 10, loss: 0.015460655093193054
step: 20, loss: 0.00021681330690626055
step: 30, loss: 6.198025948833674e-05
step: 40, loss: 0.00013566642883233726
step: 50, loss: 0.0017281502950936556
step: 60, loss: 9.522755863144994e-05
step: 70, loss: 0.00017059587116818875
step: 80, loss: 0.0005437966319732368
step: 90, loss: 0.00013141606177669019
step: 100, loss: 0.00025278920657001436

```

```

step: 110, loss: 0.00040107412496581674
step: 120, loss: 0.00016959491767920554
step: 130, loss: 7.713964441791177e-05
step: 140, loss: 7.10090171196498e-05
step: 150, loss: 0.00023205661273095757
step: 160, loss: 0.004243460483849049
step: 170, loss: 7.668539910810068e-05
step: 180, loss: 0.04893166944384575
step: 190, loss: 0.00010029390250565484
step: 200, loss: 2.1098414435982704e-05
step: 210, loss: 0.00038590573240071535
step: 220, loss: 9.522720210952684e-05
step: 230, loss: 0.0002511630882509053
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9785794813979707,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001012336288113147
step: 10, loss: 6.694886542391032e-05
step: 20, loss: 0.000234571925830096
step: 30, loss: 8.137621625792235e-05
step: 40, loss: 9.15242635528557e-05
step: 50, loss: 4.354718839749694e-05
step: 60, loss: 9.549623791826889e-05
step: 70, loss: 7.318490679608658e-05
step: 80, loss: 3.584628939279355e-05
step: 90, loss: 0.004101624712347984
step: 100, loss: 9.63111306191422e-05
step: 110, loss: 6.765270518371835e-05
step: 120, loss: 2.8739275876432657e-05
step: 130, loss: 6.223376840353012e-05
step: 140, loss: 0.00018456904217600822
step: 150, loss: 3.55685297108721e-05
step: 160, loss: 0.00010837883746717125
step: 170, loss: 0.00013235825463198125
step: 180, loss: 4.016647653770633e-05
step: 190, loss: 0.0001110169687308371
step: 200, loss: 8.069320028880611e-05
step: 210, loss: 4.65927405457478e-05
step: 220, loss: 0.0003585951926652342
step: 230, loss: 5.6799413869157434e-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=0.9844097995545658, f1=0.9799554565701558,
best_f1=0.9819819819819819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006757709779776633
step: 10, loss: 4.8857655201572925e-05
step: 20, loss: 0.0010112584568560123
step: 30, loss: 4.0371756767854095e-05
step: 40, loss: 2.44813763856655e-05
step: 50, loss: 3.478063445072621e-05
step: 60, loss: 0.04041577875614166
step: 70, loss: 0.00038051497540436685
step: 80, loss: 4.802842522622086e-05
step: 90, loss: 4.307918425183743e-05
step: 100, loss: 5.2779116231249645e-05
step: 110, loss: 6.956588185857981e-05
step: 120, loss: 0.03280755504965782
step: 130, loss: 0.000482801377074793
step: 140, loss: 0.020610706880688667
step: 150, loss: 0.002400701865553856
step: 160, loss: 0.015498901717364788
step: 170, loss: 2.079393198073376e-05
step: 180, loss: 6.642970402026549e-05
step: 190, loss: 0.00010622325498843566
step: 200, loss: 0.0016481425845995545
step: 210, loss: 0.04193241521716118
step: 220, loss: 4.424112194101326e-05
step: 230, loss: 4.056177567690611e-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=0.9844444444444443, f1=0.975609756097561,
best_f1=0.9819819819819819
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 221.06it/s]
load_f1 = 0.9854096520763187
real_f1 = 0.9799107142857142
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 [00:21, 202.45it/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: 0.6416119337081909
```

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: 0.3998073935508728
```

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

```
step: 20, loss: 0.30221429467201233
```

```
step: 30, loss: 0.3208423852920532
```

```
step: 40, loss: 0.31859487295150757
```

```
step: 50, loss: 0.46153363585472107
```

```
step: 60, loss: 0.18769939243793488
```


step: 70, loss: 0.2504480481147766
step: 80, loss: 0.19714389741420746
step: 90, loss: 0.1790952831506729
step: 100, loss: 0.23326241970062256
step: 110, loss: 0.27680256962776184
step: 120, loss: 0.0634760931134224
step: 130, loss: 0.12026692181825638
step: 140, loss: 0.2787039279937744
step: 150, loss: 0.04971971735358238
step: 160, loss: 0.06981859356164932
step: 170, loss: 0.053001582622528076
step: 180, loss: 0.03402700647711754
step: 190, loss: 0.04480140656232834
step: 200, loss: 0.06265857815742493
step: 210, loss: 0.043242309242486954
step: 220, loss: 0.08515816926956177
step: 230, loss: 0.17646725475788116
step: 240, loss: 0.02089584432542324
step: 250, loss: 0.038083478808403015
step: 260, loss: 0.17353582382202148
step: 270, loss: 0.3517371416091919
step: 280, loss: 0.05002652481198311
step: 290, loss: 0.06537044048309326
step: 300, loss: 0.06951349228620529
step: 310, loss: 0.24443672597408295
step: 320, loss: 0.09849853813648224
step: 330, loss: 0.05479412525892258
step: 340, loss: 0.27254897356033325
step: 350, loss: 0.07855039089918137
step: 360, loss: 0.10278411209583282
step: 370, loss: 0.05068252235651016
step: 380, loss: 0.10363893955945969
step: 390, loss: 0.015616098418831825
step: 400, loss: 0.0339447446167469
step: 410, loss: 0.2712126672267914
step: 420, loss: 0.041047561913728714
step: 430, loss: 0.008278047665953636
step: 440, loss: 0.008745428174734116
step: 450, loss: 0.019011186435818672
step: 460, loss: 0.006415116135030985
step: 470, loss: 0.0394926555454731
step: 480, loss: 0.14799059927463531
step: 490, loss: 0.11735604703426361
step: 500, loss: 0.024344434961676598
step: 510, loss: 0.0481744185090065
step: 520, loss: 0.06072736904025078
step: 530, loss: 0.04436061531305313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9473684210526316, f1=0.9397031539888682,
best_f1=0.9397031539888682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013407612219452858
step: 10, loss: 0.0712791234254837
step: 20, loss: 0.012946570292115211
step: 30, loss: 0.08562154322862625
step: 40, loss: 0.08005307614803314
step: 50, loss: 0.11943168193101883
step: 60, loss: 0.03768116235733032
step: 70, loss: 0.011744086630642414
step: 80, loss: 0.020468126982450485
step: 90, loss: 0.05670168623328209
step: 100, loss: 0.13171349465847015
step: 110, loss: 0.012042820453643799
step: 120, loss: 0.024846872314810753
step: 130, loss: 0.003681030124425888
step: 140, loss: 0.12605459988117218
step: 150, loss: 0.042013585567474365
step: 160, loss: 0.06696288287639618
step: 170, loss: 0.04534171521663666
step: 180, loss: 0.04240105301141739
step: 190, loss: 0.019479500129818916
step: 200, loss: 0.20920786261558533
step: 210, loss: 0.052618905901908875
step: 220, loss: 0.0014613809762522578
step: 230, loss: 0.03302972391247749
step: 240, loss: 0.07627197355031967
step: 250, loss: 0.009053406305611134
step: 260, loss: 0.03829608112573624
step: 270, loss: 0.042351093143224716
step: 280, loss: 0.0211090799421072
step: 290, loss: 0.04734700173139572
step: 300, loss: 0.020366501063108444
step: 310, loss: 0.06179925054311752
step: 320, loss: 0.025898020714521408
step: 330, loss: 0.12768881022930145
step: 340, loss: 0.1074235811829567
step: 350, loss: 0.0033743681851774454
step: 360, loss: 0.12319907546043396

```

```

step: 370, loss: 0.0063118357211351395
step: 380, loss: 0.15883223712444305
step: 390, loss: 0.0059295035898685455
step: 400, loss: 0.04041357338428497
step: 410, loss: 0.015434440225362778
step: 420, loss: 0.04997735470533371
step: 430, loss: 0.0731068104505539
step: 440, loss: 0.0341668576002121
step: 450, loss: 0.08511345833539963
step: 460, loss: 0.037210285663604736
step: 470, loss: 0.03902705758810043
step: 480, loss: 0.00799763947725296
step: 490, loss: 0.03613355755805969
step: 500, loss: 0.010285082273185253
step: 510, loss: 0.010947911068797112
step: 520, loss: 0.40259236097335815
step: 530, loss: 0.08842728286981583
/usr/local/lib/python3.7/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=0.9406819243344232, f1=0.9449112978524743,
best_f1=0.9397031539888682
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09433892369270325
step: 10, loss: 0.03867316618561745
step: 20, loss: 0.021945184096693993
step: 30, loss: 0.033816102892160416
step: 40, loss: 0.027605459094047546
step: 50, loss: 0.01934831589460373
step: 60, loss: 0.027807999402284622
step: 70, loss: 0.015899550169706345
step: 80, loss: 0.1184767559170723
step: 90, loss: 0.035993896424770355
step: 100, loss: 0.07564543187618256
step: 110, loss: 0.04264085367321968
step: 120, loss: 0.23628656566143036
step: 130, loss: 0.05444806069135666
step: 140, loss: 0.009804142639040947
step: 150, loss: 0.033254798501729965
step: 160, loss: 0.07546652853488922
step: 170, loss: 0.005884618032723665
step: 180, loss: 0.01127514336258173
step: 190, loss: 0.002190040424466133
step: 200, loss: 0.011209464631974697
step: 210, loss: 0.04200315102934837
step: 220, loss: 0.15000204741954803

```

```

step: 230, loss: 0.053733132779598236
step: 240, loss: 0.020363984629511833
step: 250, loss: 0.06553521752357483
step: 260, loss: 0.09663823246955872
step: 270, loss: 0.02090822160243988
step: 280, loss: 0.0008931165211834013
step: 290, loss: 0.024836767464876175
step: 300, loss: 0.06078706309199333
step: 310, loss: 0.06661329418420792
step: 320, loss: 0.07834522426128387
step: 330, loss: 0.019527189433574677
step: 340, loss: 0.017236359417438507
step: 350, loss: 0.15889444947242737
step: 360, loss: 0.009847206994891167
step: 370, loss: 0.06640230119228363
step: 380, loss: 0.006230618804693222
step: 390, loss: 0.0027861485723406076
step: 400, loss: 0.04985451325774193
step: 410, loss: 0.05330365151166916
step: 420, loss: 0.01028756145387888
step: 430, loss: 0.014163358137011528
step: 440, loss: 0.1085856705904007
step: 450, loss: 0.08668876439332962
step: 460, loss: 0.14597336947917938
step: 470, loss: 0.050458330661058426
step: 480, loss: 0.32135576009750366
step: 490, loss: 0.06914112716913223
step: 500, loss: 0.009176445193588734
step: 510, loss: 0.030735544860363007
step: 520, loss: 0.005675585009157658
step: 530, loss: 0.0044815754517912865
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9510945505356311, f1=0.9522484932777004,
best_f1=0.9522484932777004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.039401277899742126
step: 10, loss: 0.014602667652070522
step: 20, loss: 0.09060479700565338
step: 30, loss: 0.12510144710540771
step: 40, loss: 0.0034792269580066204

```

step: 50, loss: 0.014504448510706425
step: 60, loss: 0.05293861776590347
step: 70, loss: 0.010608015581965446
step: 80, loss: 0.011679384857416153
step: 90, loss: 0.23844832181930542
step: 100, loss: 0.0031439694575965405
step: 110, loss: 0.09304648637771606
step: 120, loss: 0.0030399567913264036
step: 130, loss: 0.07659046351909637
step: 140, loss: 0.04722336307168007
step: 150, loss: 0.006003999151289463
step: 160, loss: 0.014674190431833267
step: 170, loss: 0.06397280097007751
step: 180, loss: 0.02928623929619789
step: 190, loss: 0.06083902716636658
step: 200, loss: 0.03419411554932594
step: 210, loss: 0.009044567123055458
step: 220, loss: 0.0016749579226598144
step: 230, loss: 0.016302144154906273
step: 240, loss: 0.020823368802666664
step: 250, loss: 0.10400016605854034
step: 260, loss: 0.004276355262845755
step: 270, loss: 0.08881042897701263
step: 280, loss: 0.004381329752504826
step: 290, loss: 0.027762601152062416
step: 300, loss: 0.018728921189904213
step: 310, loss: 0.008039172738790512
step: 320, loss: 0.06920010596513748
step: 330, loss: 0.008342355489730835
step: 340, loss: 0.0009765453869476914
step: 350, loss: 0.05304276943206787
step: 360, loss: 0.06459581106901169
step: 370, loss: 0.0018835022347047925
step: 380, loss: 0.003887326456606388
step: 390, loss: 0.0023010254371911287
step: 400, loss: 0.0064156209118664265
step: 410, loss: 0.005832348484545946
step: 420, loss: 0.019622277468442917
step: 430, loss: 0.004534621257334948
step: 440, loss: 0.006620825733989477
step: 450, loss: 0.025450313463807106
step: 460, loss: 0.007653957698494196
step: 470, loss: 0.0037819717545062304
step: 480, loss: 0.012911666184663773
step: 490, loss: 0.0003729028394445777
step: 500, loss: 0.05713333934545517
step: 510, loss: 0.014101885259151459
step: 520, loss: 0.035582296550273895

```

step: 530, loss: 0.07411104440689087
/usr/local/lib/python3.7/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=0.9506398537477148, f1=0.9444444444444445,
best_f1=0.9522484932777004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007076093461364508
step: 10, loss: 0.018813803791999817
step: 20, loss: 0.004089243244379759
step: 30, loss: 0.029958834871649742
step: 40, loss: 0.014481774531304836
step: 50, loss: 0.1425570249557495
step: 60, loss: 0.05274637043476105
step: 70, loss: 0.00310884858481586
step: 80, loss: 0.0030373171903192997
step: 90, loss: 0.14134882390499115
step: 100, loss: 0.08172201365232468
step: 110, loss: 0.014824326150119305
step: 120, loss: 0.23991817235946655
step: 130, loss: 0.013163495808839798
step: 140, loss: 0.0037649699952453375
step: 150, loss: 0.020372610539197922
step: 160, loss: 0.003977909684181213
step: 170, loss: 0.050577808171510696
step: 180, loss: 0.004753547720611095
step: 190, loss: 0.0016099484637379646
step: 200, loss: 0.006701347418129444
step: 210, loss: 0.0010257090907543898
step: 220, loss: 0.0394841693341732
step: 230, loss: 0.022441890090703964
step: 240, loss: 0.009689182043075562
step: 250, loss: 0.063825324177742
step: 260, loss: 0.003591107903048396
step: 270, loss: 0.1014614850282669
step: 280, loss: 0.010311293415725231
step: 290, loss: 0.004470615182071924
step: 300, loss: 0.0421040877699852
step: 310, loss: 0.15410518646240234
step: 320, loss: 0.019104884937405586
step: 330, loss: 0.022439658641815186
step: 340, loss: 0.004814652260392904
step: 350, loss: 0.003036174923181534
step: 360, loss: 0.000506924232468009
step: 370, loss: 0.00028554126038216054
step: 380, loss: 0.00013909397239331156

```

```

step: 390, loss: 0.00017886106797959656
step: 400, loss: 0.07753156125545502
step: 410, loss: 0.03139317408204079
step: 420, loss: 0.15101182460784912
step: 430, loss: 0.03695978224277496
step: 440, loss: 0.015134595334529877
step: 450, loss: 0.016000546514987946
step: 460, loss: 0.007464439608156681
step: 470, loss: 0.0337655134499073
step: 480, loss: 0.00629771500825882
step: 490, loss: 0.01137075200676918
step: 500, loss: 0.008086426183581352
step: 510, loss: 0.0030853895004838705
step: 520, loss: 0.12772540748119354
step: 530, loss: 0.024158526211977005
/usr/local/lib/python3.7/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=0.950381679389313, f1=0.9344262295081968,
best_f1=0.9522484932777004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027724452316761017
step: 10, loss: 0.0004892938886769116
step: 20, loss: 0.0007331101805903018
step: 30, loss: 0.0006912339013069868
step: 40, loss: 0.0005382978706620634
step: 50, loss: 0.053043466061353683
step: 60, loss: 0.03616994619369507
step: 70, loss: 0.004498715046793222
step: 80, loss: 0.0022554255556315184
step: 90, loss: 0.009590713307261467
step: 100, loss: 0.007055360823869705
step: 110, loss: 0.001014422276057303
step: 120, loss: 0.07522433251142502
step: 130, loss: 0.008080597035586834
step: 140, loss: 0.0013862820342183113
step: 150, loss: 0.0005789287970401347
step: 160, loss: 0.03241344168782234
step: 170, loss: 0.0013277073157951236
step: 180, loss: 0.0016130841104313731
step: 190, loss: 0.01689537614583969
step: 200, loss: 0.0038429507985711098
step: 210, loss: 0.0020831660367548466
step: 220, loss: 0.002631682902574539
step: 230, loss: 0.0024984837509691715
step: 240, loss: 0.018790101632475853

```

```

step: 250, loss: 0.04570957273244858
step: 260, loss: 0.0039359102956950665
step: 270, loss: 0.0024965039920061827
step: 280, loss: 0.0017132461071014404
step: 290, loss: 0.0007573429029434919
step: 300, loss: 0.0008001732639968395
step: 310, loss: 0.2084895521402359
step: 320, loss: 0.013028845191001892
step: 330, loss: 0.031887155026197433
step: 340, loss: 0.003259391291067004
step: 350, loss: 0.02593752183020115
step: 360, loss: 0.050154637545347214
step: 370, loss: 0.02645745500922203
step: 380, loss: 0.0012749453308060765
step: 390, loss: 0.00733323534950614
step: 400, loss: 0.01662897691130638
step: 410, loss: 0.01751312054693699
step: 420, loss: 0.004442432429641485
step: 430, loss: 0.0013159344671294093
step: 440, loss: 0.0043485346250236034
step: 450, loss: 0.2228461056947708
step: 460, loss: 0.007197059690952301
step: 470, loss: 0.0005780891515314579
step: 480, loss: 0.0023878791835159063
step: 490, loss: 0.004840455017983913
step: 500, loss: 0.0017780085327103734
step: 510, loss: 0.23274031281471252
step: 520, loss: 0.0011778025655075908
step: 530, loss: 0.01441012043505907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.952513966480447, f1=0.9463869463869464,
best_f1=0.9463869463869464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006541553884744644
step: 10, loss: 0.005462900269776583
step: 20, loss: 0.0023556186351925135
step: 30, loss: 0.010219467803835869
step: 40, loss: 0.003934801090508699
step: 50, loss: 0.0012258286587893963
step: 60, loss: 0.0021803737618029118

```


step: 70, loss: 0.0009639133932068944
step: 80, loss: 0.00012723880354315042
step: 90, loss: 5.377924389904365e-05
step: 100, loss: 0.03752811625599861
step: 110, loss: 0.0001838414609665051
step: 120, loss: 0.0004155430942773819
step: 130, loss: 0.0015202682698145509
step: 140, loss: 0.00020873540779575706
step: 150, loss: 0.00039439124520868063
step: 160, loss: 0.00011431308666942641
step: 170, loss: 0.03310353681445122
step: 180, loss: 0.00196213461458683
step: 190, loss: 0.006806656718254089
step: 200, loss: 0.00017129092884715647
step: 210, loss: 0.0019697181414812803
step: 220, loss: 0.0008950207266025245
step: 230, loss: 8.621238521300256e-05
step: 240, loss: 0.007365671917796135
step: 250, loss: 0.009456830099225044
step: 260, loss: 0.006102856248617172
step: 270, loss: 0.0015525103081017733
step: 280, loss: 0.004535549320280552
step: 290, loss: 0.0008711523259989917
step: 300, loss: 0.0005466450820676982
step: 310, loss: 0.0004817086737602949
step: 320, loss: 0.011600786820054054
step: 330, loss: 0.008987031877040863
step: 340, loss: 0.007368181832134724
step: 350, loss: 0.000770687242038548
step: 360, loss: 0.008433341048657894
step: 370, loss: 0.02738926373422146
step: 380, loss: 0.002354402095079422
step: 390, loss: 0.027231864631175995
step: 400, loss: 0.02197570540010929
step: 410, loss: 0.0017342915525659919
step: 420, loss: 0.0975702702999115
step: 430, loss: 0.0006959334132261574
step: 440, loss: 0.015810810029506683
step: 450, loss: 0.0007622084231115878
step: 460, loss: 0.0018764790147542953
step: 470, loss: 0.002320760628208518
step: 480, loss: 0.00430306838825345
step: 490, loss: 0.02788902074098587
step: 500, loss: 0.004053922835737467
step: 510, loss: 0.0004843560454901308
step: 520, loss: 0.0020227974746376276
step: 530, loss: 0.002264037262648344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9592215013901761, f1=0.9509713228492137,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00037952649290673435
step: 10, loss: 0.0006633623270317912
step: 20, loss: 0.00037317967507988214
step: 30, loss: 0.00047214256483130157
step: 40, loss: 0.0006141993799246848
step: 50, loss: 0.010290108621120453
step: 60, loss: 0.005020620301365852
step: 70, loss: 0.002076589036732912
step: 80, loss: 0.0006970301037654281
step: 90, loss: 0.0009336420916952193
step: 100, loss: 0.0006281244568526745
step: 110, loss: 0.0010607319418340921
step: 120, loss: 0.0023823766969144344
step: 130, loss: 0.002278766129165888
step: 140, loss: 0.0009457563864998519
step: 150, loss: 0.003539043478667736
step: 160, loss: 0.0007719434215687215
step: 170, loss: 0.17169007658958435
step: 180, loss: 0.007185137365013361
step: 190, loss: 0.01436264906078577
step: 200, loss: 0.0006784616271033883
step: 210, loss: 0.0720185935497284
step: 220, loss: 0.026203934103250504
step: 230, loss: 0.05186920240521431
step: 240, loss: 0.01799992471933365
step: 250, loss: 0.00048381288070231676
step: 260, loss: 0.0022812543902546167
step: 270, loss: 0.0465206652879715
step: 280, loss: 0.0003769762988667935
step: 290, loss: 0.0027980832383036613
step: 300, loss: 0.0048010991886258125
step: 310, loss: 0.002191634848713875
step: 320, loss: 0.006029295735061169
step: 330, loss: 0.0017841863445937634
step: 340, loss: 0.027112804353237152
step: 350, loss: 0.001626664074137807
step: 360, loss: 0.02203253284096718

```

```

step: 370, loss: 0.05742073431611061
step: 380, loss: 9.056604176294059e-05
step: 390, loss: 0.0018722875975072384
step: 400, loss: 0.0008016295032575727
step: 410, loss: 0.004416937939822674
step: 420, loss: 0.00020795858290512115
step: 430, loss: 0.009329077787697315
step: 440, loss: 0.005049095954746008
step: 450, loss: 0.0004099906946066767
step: 460, loss: 0.010634913109242916
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 470, loss: 0.12068047374486923
step: 480, loss: 0.009935440495610237
step: 490, loss: 0.08316189050674438
step: 500, loss: 0.0005622307071462274
step: 510, loss: 0.00453165965154767
step: 520, loss: 0.00016759030404500663
step: 530, loss: 0.02742701955139637
/usr/local/lib/python3.7/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=0.9516658845612389, f1=0.9458823529411764,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011324434308335185
step: 10, loss: 0.003251610789448023
step: 20, loss: 0.0015014925738796592
step: 30, loss: 0.06944186985492706
step: 40, loss: 0.0012107904767617583
step: 50, loss: 0.00024412586935795844
step: 60, loss: 0.003896248061209917
step: 70, loss: 0.011683221906423569
step: 80, loss: 0.04418643191456795
step: 90, loss: 0.023588689044117928
step: 100, loss: 0.0006773031782358885
step: 110, loss: 0.005219429265707731
step: 120, loss: 0.0005219195154495537
step: 130, loss: 0.0003757785598281771
step: 140, loss: 0.0002760026545729488
step: 150, loss: 0.0005327501567080617
step: 160, loss: 0.0036082430742681026
step: 170, loss: 0.02145092375576496
step: 180, loss: 0.04001491516828537
step: 190, loss: 0.00011683553748298436
step: 200, loss: 0.0004961979575455189
step: 210, loss: 0.11106722056865692

```

```

step: 220, loss: 0.000312304706312716
step: 230, loss: 0.00024175856378860772
step: 240, loss: 0.0008222797187045217
step: 250, loss: 0.000849578995257616
step: 260, loss: 0.0018842131830751896
step: 270, loss: 0.0018861854914575815
step: 280, loss: 0.04199785366654396
step: 290, loss: 0.00031695456709712744
step: 300, loss: 0.0001090154837584123
step: 310, loss: 0.10434471070766449
step: 320, loss: 0.0046163261868059635
step: 330, loss: 0.002566338051110506
step: 340, loss: 0.05095279589295387
step: 350, loss: 0.015133647248148918
step: 360, loss: 0.004622885026037693
step: 370, loss: 0.00026351746055297554
step: 380, loss: 0.005006053484976292
step: 390, loss: 0.00011791162978624925
step: 400, loss: 0.16428200900554657
step: 410, loss: 0.014214604161679745
step: 420, loss: 0.006815549451857805
step: 430, loss: 0.006338632199913263
step: 440, loss: 0.0007006912492215633
step: 450, loss: 0.0037755481898784637
step: 460, loss: 0.00012015544052701443
step: 470, loss: 0.0002247851516585797
step: 480, loss: 0.00749631505459547
step: 490, loss: 0.026507969945669174
step: 500, loss: 0.0038420220371335745
step: 510, loss: 0.002066398039460182
step: 520, loss: 0.00968309585005045
step: 530, loss: 0.04343750327825546
/usr/local/lib/python3.7/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=0.9549046954904695, f1=0.95, best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001048819744028151
step: 10, loss: 0.00020191619114484638
step: 20, loss: 0.0030469149351119995
step: 30, loss: 0.0022312195505946875
step: 40, loss: 6.95292474119924e-05
step: 50, loss: 0.00038288754876703024
step: 60, loss: 0.007996144704520702
step: 70, loss: 0.008814753964543343
step: 80, loss: 0.0003173774457536638

```

```
step: 90, loss: 0.0009184885420836508
step: 100, loss: 0.007800242863595486
step: 110, loss: 0.007509671151638031
step: 120, loss: 0.0006508075166493654
step: 130, loss: 0.0012218985939398408
step: 140, loss: 0.00034467328805476427
step: 150, loss: 0.0017869259463623166
step: 160, loss: 0.0451190248131752
step: 170, loss: 4.207084566587582e-05
step: 180, loss: 0.0034217217471450567
step: 190, loss: 0.001118945307098329
step: 200, loss: 0.00012724613770842552
step: 210, loss: 0.006622888147830963
step: 220, loss: 6.990499241510406e-05
step: 230, loss: 3.087751247221604e-05
step: 240, loss: 6.054167170077562e-05
step: 250, loss: 0.027246618643403053
step: 260, loss: 0.0009046031045727432
step: 270, loss: 0.00011604686733335257
step: 280, loss: 0.012572798877954483
step: 290, loss: 0.0005161722656339407
step: 300, loss: 0.00025892883422784507
step: 310, loss: 0.08668594062328339
step: 320, loss: 0.0019239146495237947
step: 330, loss: 0.0017709749517962337
step: 340, loss: 5.4981177527224645e-05
step: 350, loss: 0.00017655300325714052
step: 360, loss: 8.803731907391921e-05
step: 370, loss: 0.0023757037706673145
step: 380, loss: 0.002693892689421773
step: 390, loss: 0.00045432610204443336
step: 400, loss: 0.0003372599312569946
step: 410, loss: 0.008845407515764236
step: 420, loss: 5.246156069915742e-05
step: 430, loss: 8.359427738469094e-05
step: 440, loss: 7.502428343286738e-05
step: 450, loss: 0.006100171245634556
step: 460, loss: 0.00034056761069223285
step: 470, loss: 0.017722412943840027
step: 480, loss: 0.0002445500867906958
step: 490, loss: 0.0011901288526132703
step: 500, loss: 0.0008292505517601967
step: 510, loss: 0.00022207976144272834
step: 520, loss: 0.00040534677100367844
step: 530, loss: 0.0003136330342385918
/usr/local/lib/python3.7/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=0.9566446748350613, f1=0.9496470588235294,  
best_f1=0.9509713228492137  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0010412415722385049  
step: 10, loss: 0.0022516255266964436  
step: 20, loss: 7.763513713143766e-05  
step: 30, loss: 0.0015741161769255996  
step: 40, loss: 3.8348687667166814e-05  
step: 50, loss: 0.0006306167342700064  
step: 60, loss: 3.0125831472105347e-05  
step: 70, loss: 0.0002826490963343531  
step: 80, loss: 0.0003751856565941125  
step: 90, loss: 0.00024100074369926006  
step: 100, loss: 0.00028081025811843574  
step: 110, loss: 8.748519030632451e-05  
step: 120, loss: 3.527455555740744e-05  
step: 130, loss: 1.6666657757014036e-05  
step: 140, loss: 1.680445893725846e-05  
step: 150, loss: 0.0003912505926564336  
step: 160, loss: 0.00030314375180751085  
step: 170, loss: 1.1492418707348406e-05  
step: 180, loss: 1.3299044439918362e-05  
step: 190, loss: 0.00026842745137400925  
step: 200, loss: 1.1212971003260463e-05  
step: 210, loss: 0.019193843007087708  
step: 220, loss: 0.006386490538716316  
step: 230, loss: 0.007090306840837002  
step: 240, loss: 0.00536958035081625  
step: 250, loss: 6.040426160325296e-05  
step: 260, loss: 0.00450329901650548  
step: 270, loss: 3.915546403732151e-05  
step: 280, loss: 0.0007572428439743817  
step: 290, loss: 0.0005035335198044777  
step: 300, loss: 9.004921594168991e-05  
step: 310, loss: 0.017445731908082962  
step: 320, loss: 6.324028800008819e-05  
step: 330, loss: 1.538099422759842e-05  
step: 340, loss: 0.017188500612974167  
step: 350, loss: 6.795186345698312e-05  
step: 360, loss: 0.0006420172285288572  
step: 370, loss: 0.0011711656115949154  
step: 380, loss: 0.0010790335945785046  
step: 390, loss: 0.0003816410608123988  
step: 400, loss: 0.0017644260078668594  
step: 410, loss: 0.00034420620067976415  
step: 420, loss: 2.81625434581656e-05
```

```

step: 430, loss: 0.0032058539800345898
step: 440, loss: 0.0007622899138368666
step: 450, loss: 0.0009009219356812537
step: 460, loss: 0.00031449715606868267
step: 470, loss: 0.0005235917633399367
step: 480, loss: 0.0011287906672805548
step: 490, loss: 0.00011782743968069553
step: 500, loss: 0.0006787101156078279
step: 510, loss: 0.0008971543866209686
step: 520, loss: 0.0014137179823592305
step: 530, loss: 0.0025070959236472845
/usr/local/lib/python3.7/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=0.959174096668231, f1=0.9519321394910463,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012177070311736315
step: 10, loss: 0.000933092727791518
step: 20, loss: 0.012234373018145561
step: 30, loss: 0.0004959981888532639
step: 40, loss: 0.0004614527279045433
step: 50, loss: 5.622779281111434e-05
step: 60, loss: 4.2173480323981494e-05
step: 70, loss: 0.0006504664779640734
step: 80, loss: 0.0012284251861274242
step: 90, loss: 0.003930770792067051
step: 100, loss: 0.0064978585578501225
step: 110, loss: 5.743358633480966e-05
step: 120, loss: 0.0005218470469117165
step: 130, loss: 0.000451134197646752
step: 140, loss: 0.00010080008360091597
step: 150, loss: 0.0016832883702591062
step: 160, loss: 0.0006768736056983471
step: 170, loss: 0.00018835379160009325
step: 180, loss: 2.9358094252529554e-05
step: 190, loss: 0.00010782841127365828
step: 200, loss: 6.865599425509572e-05
step: 210, loss: 4.841333793592639e-05
step: 220, loss: 2.4489432689733803e-05
step: 230, loss: 0.00013336226402316242
step: 240, loss: 0.0004692431539297104
step: 250, loss: 1.6823123587528244e-05
step: 260, loss: 3.21182596962899e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 270, loss: 0.003190840594470501

```

```

step: 280, loss: 2.106612555508036e-05
step: 290, loss: 0.00020701160246971995
step: 300, loss: 0.0006050935480743647
step: 310, loss: 4.6647175622638315e-05
step: 320, loss: 0.008366408757865429
step: 330, loss: 0.007687362842261791
step: 340, loss: 1.9695069568115287e-05
step: 350, loss: 2.449526436976157e-05
step: 360, loss: 8.949350012699142e-05
step: 370, loss: 0.00927223265171051
step: 380, loss: 9.379634138895199e-05
step: 390, loss: 0.0003641261428128928
step: 400, loss: 3.258262950112112e-05
step: 410, loss: 0.00016276445239782333
step: 420, loss: 6.897857383592054e-05
step: 430, loss: 0.0004004992370028049
step: 440, loss: 0.0017822326626628637
step: 450, loss: 0.003315868554636836
step: 460, loss: 0.000728633429389447
step: 470, loss: 0.0013660051627084613
step: 480, loss: 0.03109707124531269
step: 490, loss: 0.012079675681889057
step: 500, loss: 0.0008788632112555206
step: 510, loss: 0.024026421830058098
step: 520, loss: 0.0013140368973836303
step: 530, loss: 0.00025811977684497833
/usr/local/lib/python3.7/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=0.9573770491803278, f1=0.9527816736792893,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018985327915288508
step: 10, loss: 0.004344964399933815
step: 20, loss: 7.296576222870499e-05
step: 30, loss: 5.199574661673978e-05
step: 40, loss: 0.004793001338839531
step: 50, loss: 0.029763413593173027
step: 60, loss: 0.0007401524926535785
step: 70, loss: 5.279568722471595e-05
step: 80, loss: 0.00013343949103727937
step: 90, loss: 0.00042960268910974264
step: 100, loss: 0.0003314767964184284
step: 110, loss: 9.657708142185584e-05
step: 120, loss: 6.466911145253107e-05
step: 130, loss: 0.0009894389659166336

```



```

step: 140, loss: 7.736842235317454e-05
step: 150, loss: 0.0014439491787925363
step: 160, loss: 0.0006260535446926951
step: 170, loss: 0.0015804516151547432
step: 180, loss: 5.848823275300674e-05
step: 190, loss: 0.011245621368288994
step: 200, loss: 0.0028129282873123884
step: 210, loss: 1.8965100025525317e-05
step: 220, loss: 7.588100561406463e-05
step: 230, loss: 0.0036961666774004698
step: 240, loss: 0.0006737291114404798
step: 250, loss: 8.714264549780637e-05
step: 260, loss: 2.4429433324257843e-05
step: 270, loss: 0.002088429406285286
step: 280, loss: 7.388659287244081e-05
step: 290, loss: 0.0010929866693913937
step: 300, loss: 0.00013065754319541156
step: 310, loss: 2.642634899530094e-05
step: 320, loss: 0.0011239404557272792
step: 330, loss: 3.5816239687846974e-05
step: 340, loss: 0.001688888412900269
step: 350, loss: 2.0674495317507535e-05
step: 360, loss: 0.0009594925795681775
step: 370, loss: 0.0027397286612540483
step: 380, loss: 2.4376606234000064e-05
step: 390, loss: 0.00012995651923120022
step: 400, loss: 0.0005425116396509111
step: 410, loss: 8.951558993430808e-05
step: 420, loss: 0.0001495711476309225
step: 430, loss: 0.00014604991883970797
step: 440, loss: 1.6089214113890193e-05
step: 450, loss: 0.00011943567369598895
step: 460, loss: 0.0002464386634528637
step: 470, loss: 4.193023414700292e-05
step: 480, loss: 1.064676143869292e-05
step: 490, loss: 3.847926564048976e-05
step: 500, loss: 5.8904890465782955e-05
step: 510, loss: 9.929491352522746e-05
step: 520, loss: 2.1423547877930105e-05
step: 530, loss: 5.801140287076123e-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=0.9534450651769087, f1=0.9478584729981377,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 3.546876905602403e-05
step: 10, loss: 0.0023895218037068844
step: 20, loss: 1.36231710712309e-05
step: 30, loss: 0.00019896260346286
step: 40, loss: 0.00018306644051335752
step: 50, loss: 0.0009945448255166411
step: 60, loss: 0.00014151963114272803
step: 70, loss: 1.950874866452068e-05
step: 80, loss: 0.0008514419896528125
step: 90, loss: 1.2285901902941987e-05
step: 100, loss: 0.0001027335092658177
step: 110, loss: 0.008081716485321522
step: 120, loss: 1.2569031241582707e-05
step: 130, loss: 2.7340849555912428e-05
step: 140, loss: 0.0010642148554325104
step: 150, loss: 0.00159894535318017
step: 160, loss: 6.095876597100869e-05
step: 170, loss: 0.0028380875010043383
step: 180, loss: 0.00043113998253829777
step: 190, loss: 0.00010603887494653463
step: 200, loss: 1.8446937247063033e-05
step: 210, loss: 0.00014092156197875738
step: 220, loss: 2.0186451365589164e-05
step: 230, loss: 4.405486834002659e-05
step: 240, loss: 2.853863588825334e-05
step: 250, loss: 0.002211369341239333
step: 260, loss: 0.0012222876539453864
step: 270, loss: 6.0560643760254607e-05
step: 280, loss: 6.0314407164696604e-05
step: 290, loss: 0.0001509936701040715
step: 300, loss: 0.00014444059343077242
step: 310, loss: 0.0003466281632427126
step: 320, loss: 3.890058724209666e-05
step: 330, loss: 9.973827400244772e-05
step: 340, loss: 8.426132262684405e-05
step: 350, loss: 1.7112701243604533e-05
step: 360, loss: 2.3858085114625283e-05
step: 370, loss: 0.0011131155770272017
step: 380, loss: 7.495356840081513e-05
step: 390, loss: 0.0006227802368812263
step: 400, loss: 0.0012859755661338568
step: 410, loss: 0.00011050476314267144
step: 420, loss: 3.408246993785724e-05
step: 430, loss: 0.00022172476747073233
step: 440, loss: 0.01555478386580944
step: 450, loss: 1.5061058547871653e-05
step: 460, loss: 0.03528565540909767
step: 470, loss: 8.72084274305962e-06

```

step: 480, loss: 4.32144952355884e-05
step: 490, loss: 1.993524892895948e-05
step: 500, loss: 0.000741795462090522
step: 510, loss: 0.029398426413536072
step: 520, loss: 0.0006212505395524204
step: 530, loss: 0.0007688148180022836
/usr/local/lib/python3.7/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=0.9540937056318032, f1=0.9490566037735849,
best_f1=0.9509713228492137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.957619113265537e-06
step: 10, loss: 0.0010519211646169424
step: 20, loss: 0.0019466434605419636
step: 30, loss: 0.018572794273495674
step: 40, loss: 1.3213433703640476e-05
step: 50, loss: 2.5553164960001595e-05
step: 60, loss: 3.830757486866787e-05
step: 70, loss: 1.3101582226227038e-05
step: 80, loss: 4.2119037971133366e-05
step: 90, loss: 1.7615777323953807e-05
step: 100, loss: 9.39138772082515e-06
step: 110, loss: 0.0014966420130804181
step: 120, loss: 9.771350960363634e-06
step: 130, loss: 9.698910434963182e-05
step: 140, loss: 4.843928763875738e-05
step: 150, loss: 8.687314220878761e-06
step: 160, loss: 9.152971870207693e-06
step: 170, loss: 9.037463314598426e-06
step: 180, loss: 9.874453826341778e-05
step: 190, loss: 0.0011226562783122063
step: 200, loss: 0.00031251125619746745
step: 210, loss: 7.366607314907014e-05
step: 220, loss: 1.988730400626082e-05
step: 230, loss: 0.00022863087360747159
step: 240, loss: 2.7870069970958866e-05
step: 250, loss: 7.260347774717957e-05
step: 260, loss: 6.318062332866248e-06
step: 270, loss: 3.90104905818589e-05
step: 280, loss: 9.875605428533163e-06
step: 290, loss: 0.0015415253583341837
step: 300, loss: 1.0326385563530494e-05
step: 310, loss: 0.021950501948595047
step: 320, loss: 1.2554030945466366e-05
step: 330, loss: 1.2330546269367915e-05

```

```

step: 340, loss: 1.7236228813999332e-05
step: 350, loss: 0.002097242046147585
step: 360, loss: 1.1026739230146632e-05
step: 370, loss: 0.0001562518737046048
step: 380, loss: 3.9117876440286636e-05
step: 390, loss: 2.632040013850201e-05
step: 400, loss: 0.013835415244102478
step: 410, loss: 1.951497142727021e-05
step: 420, loss: 1.2725357009912841e-05
step: 430, loss: 6.251007107493933e-06
step: 440, loss: 1.1246466783632059e-05
step: 450, loss: 3.961880793212913e-05
step: 460, loss: 0.0035428395494818687
step: 470, loss: 1.1950431144214235e-05
step: 480, loss: 0.0024323395919054747
step: 490, loss: 0.000605289766099304
step: 500, loss: 0.0005846338463015854
step: 510, loss: 7.715030733379535e-06
step: 520, loss: 0.00011150109639856964
step: 530, loss: 4.2858959204750136e-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=0.9536862003780718, f1=0.9467232437529468,
best_f1=0.9509713228492137
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.

```

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
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, 282.36it/s]
load_f1 = 0.9564419990829894
real_f1 = 0.9519010535959689
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, 205.07it/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        : 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: 0.5084016919136047

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

step: 10, loss: 0.4522448778152466

step: 20, loss: 0.4580802619457245

step: 30, loss: 0.3859758973121643

step: 40, loss: 0.34943047165870667

step: 50, loss: 0.4149608612060547

step: 60, loss: 0.4552666246891022

step: 70, loss: 0.3262123465538025

```

step: 80, loss: 0.3905588686466217
step: 90, loss: 0.2547845244407654
step: 100, loss: 0.23524188995361328
step: 110, loss: 0.2642418444156647
step: 120, loss: 0.4133983850479126
step: 130, loss: 0.23770476877689362
step: 140, loss: 0.41747283935546875
step: 150, loss: 0.2836551368236542
step: 160, loss: 0.507611870765686
step: 170, loss: 0.1767638474702835
step: 180, loss: 0.3026776611804962
step: 190, loss: 0.38983020186424255
step: 200, loss: 0.16321700811386108
step: 210, loss: 0.36047813296318054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5773195876288658, f1=0.5507246376811593,
best_f1=0.5507246376811593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2950092554092407
step: 10, loss: 0.105188749730587
step: 20, loss: 0.3733397126197815
step: 30, loss: 0.19815713167190552
step: 40, loss: 0.5287597179412842
step: 50, loss: 0.10505364835262299
step: 60, loss: 0.22689791023731232
step: 70, loss: 0.15314318239688873
step: 80, loss: 0.2458643615245819
step: 90, loss: 0.18194283545017242
step: 100, loss: 0.31685343384742737
step: 110, loss: 0.298540323972702
step: 120, loss: 0.054822713136672974
step: 130, loss: 0.1618717610836029
step: 140, loss: 0.14276033639907837
step: 150, loss: 0.2630503177642822
step: 160, loss: 0.07141533493995667
step: 170, loss: 0.26211121678352356
step: 180, loss: 0.18885649740695953
step: 190, loss: 0.30930018424987793
step: 200, loss: 0.12364929169416428
step: 210, loss: 0.12698280811309814

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6515151515151516, f1=0.6954813359528488,
best_f1=0.6954813359528488
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0777992233633995
step: 10, loss: 0.05209881067276001
step: 20, loss: 0.20613974332809448
step: 30, loss: 0.08180832117795944
step: 40, loss: 0.31935548782348633
step: 50, loss: 0.1302485316991806
step: 60, loss: 0.26090946793556213
step: 70, loss: 0.1042882651090622
step: 80, loss: 0.13862884044647217
step: 90, loss: 0.0940377414226532
step: 100, loss: 0.3166659474372864
step: 110, loss: 0.1570553481578827
step: 120, loss: 0.1647801399230957
step: 130, loss: 0.19306470453739166
step: 140, loss: 0.12461596727371216
step: 150, loss: 0.1662427932024002
step: 160, loss: 0.19724008440971375
step: 170, loss: 0.14014525711536407
step: 180, loss: 0.15803560614585876
step: 190, loss: 0.05889889597892761
step: 200, loss: 0.08104698359966278
step: 210, loss: 0.19024311006069183
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6556169429097606, f1=0.6892655367231638,
best_f1=0.6892655367231638
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061463363468647
step: 10, loss: 0.06388641893863678

```



```

step: 20, loss: 0.075662761926651
step: 30, loss: 0.06369919329881668
step: 40, loss: 0.08213350921869278
step: 50, loss: 0.16902640461921692
step: 60, loss: 0.2454979419708252
step: 70, loss: 0.06966539472341537
step: 80, loss: 0.20526401698589325
step: 90, loss: 0.09837714582681656
step: 100, loss: 0.13045132160186768
step: 110, loss: 0.32650652527809143
step: 120, loss: 0.13548840582370758
step: 130, loss: 0.21102374792099
step: 140, loss: 0.24220700562000275
step: 150, loss: 0.11770174652338028
step: 160, loss: 0.404962420463562
step: 170, loss: 0.05244232714176178
step: 180, loss: 0.062107574194669724
step: 190, loss: 0.0597551055252552
step: 200, loss: 0.06300333142280579
step: 210, loss: 0.25995752215385437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7053941908713692, f1=0.7245762711864407,
best_f1=0.7245762711864407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1303221881389618
step: 10, loss: 0.15996330976486206
step: 20, loss: 0.1612599790096283
step: 30, loss: 0.028348596766591072
step: 40, loss: 0.07197661697864532
step: 50, loss: 0.14249448478221893
step: 60, loss: 0.1473141461610794
step: 70, loss: 0.21155479550361633
step: 80, loss: 0.1098545491695404
step: 90, loss: 0.044512439519166946
step: 100, loss: 0.029145779088139534
step: 110, loss: 0.0923776626586914
step: 120, loss: 0.15894681215286255
step: 130, loss: 0.12409524619579315
step: 140, loss: 0.11538632214069366
step: 150, loss: 0.17745409905910492

```

```

step: 160, loss: 0.029211703687906265
step: 170, loss: 0.08586284518241882
step: 180, loss: 0.11408951878547668
step: 190, loss: 0.08071276545524597
step: 200, loss: 0.06355485320091248
step: 210, loss: 0.0753389298915863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7224334600760457, f1=0.7436399217221135,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010046208277344704
step: 10, loss: 0.25362011790275574
step: 20, loss: 0.03217046707868576
step: 30, loss: 0.057869039475917816
step: 40, loss: 0.0931616947054863
step: 50, loss: 0.07850459963083267
step: 60, loss: 0.15360106527805328
step: 70, loss: 0.04998095706105232
step: 80, loss: 0.10516773164272308
step: 90, loss: 0.05401385575532913
step: 100, loss: 0.10062556713819504
step: 110, loss: 0.05788445100188255
step: 120, loss: 0.02631068043410778
step: 130, loss: 0.047094594687223434
step: 140, loss: 0.11448705941438675
step: 150, loss: 0.046357762068510056
step: 160, loss: 0.11872771382331848
step: 170, loss: 0.0776476114988327
step: 180, loss: 0.06316718459129333
step: 190, loss: 0.06683432310819626
step: 200, loss: 0.07885011285543442
step: 210, loss: 0.07900456339120865
/usr/local/lib/python3.7/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=0.7040618955512573, f1=0.7211155378486056,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.13792093098163605
step: 10, loss: 0.018365276977419853
step: 20, loss: 0.03174867853522301
step: 30, loss: 0.19012793898582458
step: 40, loss: 0.017055189236998558
step: 50, loss: 0.07639855891466141
step: 60, loss: 0.1299641728401184
step: 70, loss: 0.07514309883117676
step: 80, loss: 0.038749828934669495
step: 90, loss: 0.21637018024921417
step: 100, loss: 0.10507508367300034
step: 110, loss: 0.11772589385509491
step: 120, loss: 0.053733136504888535
step: 130, loss: 0.2959892153739929
step: 140, loss: 0.05625898018479347
step: 150, loss: 0.036971352994441986
step: 160, loss: 0.1626890003681183
step: 170, loss: 0.09585339576005936
step: 180, loss: 0.03815601393580437
step: 190, loss: 0.09697234630584717
step: 200, loss: 0.016057586297392845
step: 210, loss: 0.12824587523937225
/usr/local/lib/python3.7/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=0.7024952015355086, f1=0.7329434697855752,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16764454543590546
step: 10, loss: 0.012253145687282085
step: 20, loss: 0.03864273801445961
step: 30, loss: 0.02430952899158001
step: 40, loss: 0.013664978556334972
step: 50, loss: 0.0077062444761395454
step: 60, loss: 0.12291653454303741
step: 70, loss: 0.07621971517801285
step: 80, loss: 0.039080195128917694
step: 90, loss: 0.16346779465675354
step: 100, loss: 0.17804092168807983
step: 110, loss: 0.03470546379685402
step: 120, loss: 0.12720553576946259
step: 130, loss: 0.021067341789603233
step: 140, loss: 0.036419495940208435
step: 150, loss: 0.07069873064756393
step: 160, loss: 0.13243135809898376
step: 170, loss: 0.08256031572818756

```

```

step: 180, loss: 0.08876069635152817
step: 190, loss: 0.004088695626705885
step: 200, loss: 0.08250556886196136
step: 210, loss: 0.09231464564800262
/usr/local/lib/python3.7/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=0.7027027027027026, f1=0.7207920792079208,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029523735865950584
step: 10, loss: 0.010329102165997028
step: 20, loss: 0.06355837732553482
step: 30, loss: 0.012337001971900463
step: 40, loss: 0.08125419914722443
step: 50, loss: 0.3039582073688507
step: 60, loss: 0.019871709868311882
step: 70, loss: 0.03133382648229599
step: 80, loss: 0.01849127560853958
step: 90, loss: 0.16850736737251282
step: 100, loss: 0.06758688390254974
step: 110, loss: 0.0718047246336937
step: 120, loss: 0.23963534832000732
step: 130, loss: 0.0557267852127552
step: 140, loss: 0.028927017003297806
step: 150, loss: 0.020975107327103615
step: 160, loss: 0.15895895659923553
step: 170, loss: 0.16529062390327454
step: 180, loss: 0.06198849901556969
step: 190, loss: 0.011625401675701141
step: 200, loss: 0.010591503232717514
step: 210, loss: 0.1079094409942627
/usr/local/lib/python3.7/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=0.6987951807228915, f1=0.7419354838709677,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03242966905236244
step: 10, loss: 0.08806926012039185
step: 20, loss: 0.010161228477954865
step: 30, loss: 0.009102283976972103
step: 40, loss: 0.08000656217336655
step: 50, loss: 0.005375029519200325

```

```

step: 60, loss: 0.0035057025961577892
step: 70, loss: 0.034070536494255066
step: 80, loss: 0.047596219927072525
step: 90, loss: 0.0038516870699822903
step: 100, loss: 0.053338583558797836
step: 110, loss: 0.004756992217153311
step: 120, loss: 0.06930653005838394
step: 130, loss: 0.017631273716688156
step: 140, loss: 0.03449859842658043
step: 150, loss: 0.11987820267677307
step: 160, loss: 0.06370270997285843
step: 170, loss: 0.06577356904745102
step: 180, loss: 0.007700735237449408
step: 190, loss: 0.020078623667359352
step: 200, loss: 0.19189518690109253
step: 210, loss: 0.10635427385568619
/usr/local/lib/python3.7/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=0.6968503937007874, f1=0.7294589178356714,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.050609506666666031
step: 10, loss: 0.007381280418485403
step: 20, loss: 0.03375156223773956
step: 30, loss: 0.0331793837249279
step: 40, loss: 0.020168906077742577
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.004691118840128183
step: 60, loss: 0.00897076167166233
step: 70, loss: 0.03312891349196434
step: 80, loss: 0.05988967791199684
step: 90, loss: 0.1782074123620987
step: 100, loss: 0.04828345775604248
step: 110, loss: 0.010477096773684025
step: 120, loss: 0.006485233549028635
step: 130, loss: 0.0019869981333613396
step: 140, loss: 0.014694830402731895
step: 150, loss: 0.036800019443035126
step: 160, loss: 0.0048738885670900345
step: 170, loss: 0.019360385835170746
step: 180, loss: 0.012206695042550564
step: 190, loss: 0.07556788623332977
step: 200, loss: 0.014821868389844894
step: 210, loss: 0.03856508061289787
/usr/local/lib/python3.7/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=0.7072691552062869, f1=0.7416173570019724,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02724555693566799
step: 10, loss: 0.03821665793657303
step: 20, loss: 0.049136243760585785
step: 30, loss: 0.01082482747733593
step: 40, loss: 0.06785507500171661
step: 50, loss: 0.04377821460366249
step: 60, loss: 0.003843148471787572
step: 70, loss: 0.012699984945356846
step: 80, loss: 0.030178284272551537
step: 90, loss: 0.016438743099570274
step: 100, loss: 0.0007561451639048755
step: 110, loss: 0.03047022596001625
step: 120, loss: 0.0031529313419014215
step: 130, loss: 0.12159867584705353
step: 140, loss: 0.03158673271536827
step: 150, loss: 0.0008315122104249895
step: 160, loss: 0.013208203949034214
step: 170, loss: 0.06410981714725494
step: 180, loss: 0.020936554297804832
step: 190, loss: 0.006173174828290939
step: 200, loss: 0.012483973987400532
step: 210, loss: 0.004550869110971689
/usr/local/lib/python3.7/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=0.6980392156862745, f1=0.732, best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015013207448646426
step: 10, loss: 0.009671244770288467
step: 20, loss: 0.009800911881029606
step: 30, loss: 0.0033525358885526657
step: 40, loss: 0.045221056789159775
step: 50, loss: 0.11610928922891617
step: 60, loss: 0.003200049512088299
step: 70, loss: 0.20079150795936584
step: 80, loss: 0.17127539217472076
step: 90, loss: 0.0057207378558814526
step: 100, loss: 0.04270019754767418
step: 110, loss: 0.021403970196843147

```

```

step: 120, loss: 0.035235557705163956
step: 130, loss: 0.002074030227959156
step: 140, loss: 0.02454616129398346
step: 150, loss: 0.057322222739458084
step: 160, loss: 0.022534718737006187
step: 170, loss: 0.05024634674191475
step: 180, loss: 0.001466565765440464
step: 190, loss: 0.032982971519231796
step: 200, loss: 0.029711786657571793
step: 210, loss: 0.016701823100447655
/usr/local/lib/python3.7/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=0.6902985074626866, f1=0.7279411764705882,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004722613841295242
step: 10, loss: 0.025342833250761032
step: 20, loss: 0.01278652437031269
step: 30, loss: 0.3336058557033539
step: 40, loss: 0.04108027368783951
step: 50, loss: 0.06177449971437454
step: 60, loss: 0.02539440616965294
step: 70, loss: 0.014717425219714642
step: 80, loss: 0.03150574117898941
step: 90, loss: 0.002599452156573534
step: 100, loss: 0.015423834323883057
step: 110, loss: 0.022400733083486557
step: 120, loss: 0.0017638918943703175
step: 130, loss: 0.014832987450063229
step: 140, loss: 0.1017497330904007
step: 150, loss: 0.07585495710372925
step: 160, loss: 0.02152029611170292
step: 170, loss: 0.029724614694714546
step: 180, loss: 0.0028133136220276356
step: 190, loss: 0.05439244955778122
step: 200, loss: 0.012133206240832806
step: 210, loss: 0.0032762850169092417
/usr/local/lib/python3.7/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=0.6994328922495273, f1=0.7418738049713194,
best_f1=0.7436399217221135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.010769650340080261
step: 10, loss: 0.002883884822949767
step: 20, loss: 0.06004597246646881
step: 30, loss: 0.03267301246523857
step: 40, loss: 0.0026204471942037344
step: 50, loss: 0.0006927368231117725
step: 60, loss: 0.05290141701698303
step: 70, loss: 0.06187431514263153
step: 80, loss: 0.044078897684812546
step: 90, loss: 0.010157686658203602
step: 100, loss: 0.007025057449936867
step: 110, loss: 0.08015983551740646
step: 120, loss: 0.029029566794633865
step: 130, loss: 0.008824185468256474
step: 140, loss: 0.008252724073827267
step: 150, loss: 0.005280749872326851
step: 160, loss: 0.1021801084280014
step: 170, loss: 0.03095700964331627
step: 180, loss: 0.00787657406181097
step: 190, loss: 0.09520558267831802
step: 200, loss: 0.002911698305979371
step: 210, loss: 0.022551028057932854
/usr/local/lib/python3.7/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=0.7040618955512573, f1=0.745631067961165,
best_f1=0.7436399217221135
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, 519.77it/s]
load_f1 = 0.7050359712230215
real_f1 = 0.7058823529411765
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, 203.78it/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: 0.44400903582572937

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

step: 10, loss: 0.40050193667411804

step: 20, loss: 0.2592300772666931

step: 30, loss: 0.4084942042827606

step: 40, loss: 0.2511971890926361

step: 50, loss: 0.3173554837703705

```

step: 60, loss: 0.4453095495700836
step: 70, loss: 0.4009024500846863
step: 80, loss: 0.19028718769550323
step: 90, loss: 0.2970103919506073
step: 100, loss: 0.46441641449928284
step: 110, loss: 0.237317755818367
step: 120, loss: 0.32640090584754944
step: 130, loss: 0.3203135132789612
step: 140, loss: 0.206299290060997
step: 150, loss: 0.3120497465133667
step: 160, loss: 0.23798581957817078
step: 170, loss: 0.37558630108833313
step: 180, loss: 0.1768764853477478
step: 190, loss: 0.16034866869449615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.39628082513809204
step: 10, loss: 0.3077264428138733
step: 20, loss: 0.5987556576728821
step: 30, loss: 0.23513109982013702
step: 40, loss: 0.5609724521636963
step: 50, loss: 0.3197091519832611
step: 60, loss: 0.4429733157157898
step: 70, loss: 0.3000420033931732
step: 80, loss: 0.14499720931053162
step: 90, loss: 0.3326724171638489
step: 100, loss: 0.24453218281269073
step: 110, loss: 0.3604312539100647
step: 120, loss: 0.22712048888206482
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.49493029713630676
step: 140, loss: 0.32876962423324585
step: 150, loss: 0.32162633538246155
step: 160, loss: 0.3038584887981415
step: 170, loss: 0.24214443564414978
step: 180, loss: 0.17122116684913635
step: 190, loss: 0.24366821348667145
/usr/local/lib/python3.7/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: 0.3910350203514099
step: 10, loss: 0.38291656970977783
step: 20, loss: 0.46574386954307556
step: 30, loss: 0.32518190145492554
step: 40, loss: 0.08658891916275024
step: 50, loss: 0.38000723719596863
step: 60, loss: 0.16181284189224243
step: 70, loss: 0.38140302896499634
step: 80, loss: 0.3224620521068573
step: 90, loss: 0.3590298593044281
step: 100, loss: 0.5416388511657715
step: 110, loss: 0.6752495765686035
step: 120, loss: 0.3712814152240753
step: 130, loss: 0.15526333451271057
step: 140, loss: 0.3745768368244171
step: 150, loss: 0.3099367916584015
step: 160, loss: 0.6104922890663147
step: 170, loss: 0.44967299699783325
step: 180, loss: 0.3758825361728668
step: 190, loss: 0.16253794729709625
/usr/local/lib/python3.7/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: 0.233431875705719
step: 10, loss: 0.2369903326034546
step: 20, loss: 0.3100631833076477
step: 30, loss: 0.23857086896896362
step: 40, loss: 0.5465258359909058
step: 50, loss: 0.2405664622783661
step: 60, loss: 0.38337624073028564
step: 70, loss: 0.3200809359550476
step: 80, loss: 0.23954902589321136
step: 90, loss: 0.1750234067440033
step: 100, loss: 0.3120041489601135
step: 110, loss: 0.38507604598999023
step: 120, loss: 0.24728699028491974

```

```

step: 130, loss: 0.4632089138031006
step: 140, loss: 0.30900999903678894
step: 150, loss: 0.24097350239753723
step: 160, loss: 0.306052029132843
step: 170, loss: 0.44259580969810486
step: 180, loss: 0.3881928622722626
step: 190, loss: 0.16604790091514587
/usr/local/lib/python3.7/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: 0.39376041293144226
step: 10, loss: 0.3904399275779724
step: 20, loss: 0.18159069120883942
step: 30, loss: 0.09699142724275589
step: 40, loss: 0.30379557609558105
step: 50, loss: 0.5474538207054138
step: 60, loss: 0.2541816830635071
step: 70, loss: 0.3890571892261505
step: 80, loss: 0.40637916326522827
step: 90, loss: 0.3394075334072113
step: 100, loss: 0.44564053416252136
step: 110, loss: 0.3959384560585022
step: 120, loss: 0.22665786743164062
step: 130, loss: 0.5728938579559326
step: 140, loss: 0.40268754959106445
step: 150, loss: 0.32704851031303406
step: 160, loss: 0.17874251306056976
step: 170, loss: 0.3759610950946808
step: 180, loss: 0.25252044200897217
step: 190, loss: 0.31415411829948425
/usr/local/lib/python3.7/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: 0.3243970274925232
step: 10, loss: 0.2434556782245636
step: 20, loss: 0.3106021583080292
step: 30, loss: 0.4676821231842041
step: 40, loss: 0.24729172885417938

```

```

step: 50, loss: 0.3090687692165375
step: 60, loss: 0.43616247177124023
step: 70, loss: 0.33490419387817383
step: 80, loss: 0.3045322299003601
step: 90, loss: 0.23890355229377747
step: 100, loss: 0.4995204210281372
step: 110, loss: 0.23255163431167603
step: 120, loss: 0.4652700424194336
step: 130, loss: 0.5398785471916199
step: 140, loss: 0.17642205953598022
step: 150, loss: 0.38131046295166016
step: 160, loss: 0.3799751400947571
step: 170, loss: 0.3841361105442047
step: 180, loss: 0.17289182543754578
step: 190, loss: 0.30836841464042664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3834808259587021, f1=0.3902439024390243,
best_f1=0.3902439024390243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.310452938079834
step: 10, loss: 0.2901516854763031
step: 20, loss: 0.319454163312912
step: 30, loss: 0.23164142668247223
step: 40, loss: 0.3093459904193878
step: 50, loss: 0.07327111065387726
step: 60, loss: 0.15454527735710144
step: 70, loss: 0.16280953586101532
step: 80, loss: 0.23426268994808197
step: 90, loss: 0.2380935251712799
step: 100, loss: 0.5670733451843262
step: 110, loss: 0.4357733130455017
step: 120, loss: 0.39256635308265686
step: 130, loss: 0.2986449599266052
step: 140, loss: 0.2214173972606659
step: 150, loss: 0.2873055040836334
step: 160, loss: 0.3224036395549774
step: 170, loss: 0.3221306800842285
step: 180, loss: 0.2639542818069458
step: 190, loss: 0.259537935256958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5481239804241436, f1=0.5271565495207667,
best_f1=0.5271565495207667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21192246675491333
step: 10, loss: 0.4255911409854889
step: 20, loss: 0.11952206492424011
step: 30, loss: 0.2752978503704071
step: 40, loss: 0.1730402261018753
step: 50, loss: 0.2918839454650879
step: 60, loss: 0.23186510801315308
step: 70, loss: 0.14821310341358185
step: 80, loss: 0.4176969528198242
step: 90, loss: 0.1440662294626236
step: 100, loss: 0.16110368072986603
step: 110, loss: 0.17980539798736572
step: 120, loss: 0.08224303275346756
step: 130, loss: 0.07596110552549362
step: 140, loss: 0.05899960547685623
step: 150, loss: 0.16734857857227325
step: 160, loss: 0.04299135506153107
step: 170, loss: 0.04221828654408455
step: 180, loss: 0.25450271368026733
step: 190, loss: 0.08285468816757202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.783289817232376, f1=0.7807486631016043,
best_f1=0.7807486631016043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024454018101096153
step: 10, loss: 0.04703658074140549
step: 20, loss: 0.030640235170722008
step: 30, loss: 0.029485510662198067
step: 40, loss: 0.4022027850151062

```

```

step: 50, loss: 0.22552675008773804
step: 60, loss: 0.21692639589309692
step: 70, loss: 0.031053099781274796
step: 80, loss: 0.3575398027896881
step: 90, loss: 0.11614838242530823
step: 100, loss: 0.05879073962569237
step: 110, loss: 0.07061675935983658
step: 120, loss: 0.11500933766365051
step: 130, loss: 0.11036735773086548
step: 140, loss: 0.25270378589630127
step: 150, loss: 0.04647276550531387
step: 160, loss: 0.05107909068465233
step: 170, loss: 0.24051600694656372
step: 180, loss: 0.2194875031709671
step: 190, loss: 0.011262410320341587
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8181818181818181, f1=0.8131868131868132,
best_f1=0.8131868131868132
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047384925186634064
step: 10, loss: 0.11472787708044052
step: 20, loss: 0.08897161483764648
step: 30, loss: 0.2786914110183716
step: 40, loss: 0.07479844242334366
step: 50, loss: 0.07322812080383301
step: 60, loss: 0.05605601891875267
step: 70, loss: 0.03376714885234833
step: 80, loss: 0.2957766652107239
step: 90, loss: 0.0517171174287796
step: 100, loss: 0.07547652721405029
step: 110, loss: 0.05312810465693474
step: 120, loss: 0.007966847158968449
step: 130, loss: 0.03171553090214729
step: 140, loss: 0.12185245007276535
step: 150, loss: 0.07317420840263367
step: 160, loss: 0.03251629322767258
step: 170, loss: 0.0338105708360672
step: 180, loss: 0.05933894217014313
step: 190, loss: 0.01454774197191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8302872062663187, f1=0.8263157894736842,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051755309104919434
step: 10, loss: 0.005014659371227026
step: 20, loss: 0.11583251506090164
step: 30, loss: 0.14196811616420746
step: 40, loss: 0.13188405334949493
step: 50, loss: 0.08113392442464828
step: 60, loss: 0.035098444670438766
step: 70, loss: 0.01921425573527813
step: 80, loss: 0.16921457648277283
step: 90, loss: 0.02126849815249443
step: 100, loss: 0.1497737616300583
step: 110, loss: 0.20058481395244598
step: 120, loss: 0.04119348153471947
step: 130, loss: 0.21463549137115479
step: 140, loss: 0.12649250030517578
step: 150, loss: 0.009738217107951641
step: 160, loss: 0.08509241044521332
step: 170, loss: 0.08651269972324371
step: 180, loss: 0.1033577173948288
step: 190, loss: 0.09891064465045929
/usr/local/lib/python3.7/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=0.8241758241758241, f1=0.8412256267409471,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01291116513311863
step: 10, loss: 0.03147685527801514
step: 20, loss: 0.055306095629930496
step: 30, loss: 0.022493133321404457
step: 40, loss: 0.0045177750289440155
step: 50, loss: 0.025556568056344986
step: 60, loss: 0.008676769211888313
step: 70, loss: 0.036987509578466415
step: 80, loss: 0.11727659404277802

```

```

step: 90, loss: 0.11773605644702911
step: 100, loss: 0.16312019526958466
step: 110, loss: 0.26030218601226807
step: 120, loss: 0.029831957072019577
step: 130, loss: 0.006645473651587963
step: 140, loss: 0.04435999318957329
step: 150, loss: 0.23170256614685059
step: 160, loss: 0.04001681134104729
step: 170, loss: 0.1144346222281456
step: 180, loss: 0.0064871045760810375
step: 190, loss: 0.1458522528409958
/usr/local/lib/python3.7/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=0.8282828282828283, f1=0.837696335078534,
best_f1=0.8263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26621347665786743
step: 10, loss: 0.007620898075401783
step: 20, loss: 0.04811641201376915
step: 30, loss: 0.19849567115306854
step: 40, loss: 0.04816840589046478
step: 50, loss: 0.009833008050918579
step: 60, loss: 0.013347064144909382
step: 70, loss: 0.045982856303453445
step: 80, loss: 0.037962980568408966
step: 90, loss: 0.012744924053549767
step: 100, loss: 0.003083678660914302
step: 110, loss: 0.039653196930885315
step: 120, loss: 0.01535785011947155
step: 130, loss: 0.2200203388929367
step: 140, loss: 0.200198695063591
step: 150, loss: 0.016282130032777786
step: 160, loss: 0.12984506785869598
step: 170, loss: 0.015028499066829681
step: 180, loss: 0.01357433944940567
step: 190, loss: 0.2826877236366272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8451443569553805, f1=0.8488063660477454,
best_f1=0.8488063660477454

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09299656748771667
step: 10, loss: 0.04212379828095436
step: 20, loss: 0.023511473089456558
step: 30, loss: 0.005056400317698717
step: 40, loss: 0.009199189953505993
step: 50, loss: 0.027181100100278854
step: 60, loss: 0.09871526807546616
step: 70, loss: 0.012155802920460701
step: 80, loss: 0.00978860817849636
step: 90, loss: 0.019880836829543114
step: 100, loss: 0.13427887856960297
step: 110, loss: 0.026481609791517258
step: 120, loss: 0.05546882376074791
step: 130, loss: 0.015694938600063324
step: 140, loss: 0.0038481424562633038
step: 150, loss: 0.0690915584564209
step: 160, loss: 0.011076436378061771
step: 170, loss: 0.007560707163065672
step: 180, loss: 0.034713517874479294
step: 190, loss: 0.030989404767751694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8275862068965517, f1=0.8509485094850948,
best_f1=0.8488063660477454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01239851862192154
step: 10, loss: 0.012160340324044228
step: 20, loss: 0.07171554118394852
step: 30, loss: 0.013907214626669884
step: 40, loss: 0.044259510934352875
step: 50, loss: 0.02622598223388195
step: 60, loss: 0.005583223421126604
step: 70, loss: 0.005522316321730614
step: 80, loss: 0.0025521826464682817
step: 90, loss: 0.2114802449941635
step: 100, loss: 0.21442139148712158
step: 110, loss: 0.019903037697076797
step: 120, loss: 0.013893048278987408
step: 130, loss: 0.0049231769517064095
step: 140, loss: 0.016199393197894096
step: 150, loss: 0.006105971988290548
step: 160, loss: 0.006489332299679518

```

```

step: 170, loss: 0.031593143939971924
step: 180, loss: 0.015155946835875511
step: 190, loss: 0.23659920692443848
/usr/local/lib/python3.7/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=0.8272251308900523, f1=0.8390501319261214,
best_f1=0.8488063660477454
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 [00:07, 270.64it/s]
load_f1 = 0.8033240997229917
real_f1 = 0.7967479674796748
733it [00:00, 3194.68it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 203.34it/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: 0.4992135167121887
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4661584198474884
step: 20, loss: 0.28633105754852295
step: 30, loss: 0.38130590319633484
step: 40, loss: 0.5721332430839539
step: 50, loss: 0.3731003999710083
step: 60, loss: 0.5844624042510986
step: 70, loss: 0.2955228388309479
step: 80, loss: 0.2483188658952713
step: 90, loss: 0.23503883183002472
step: 100, loss: 0.13922812044620514
step: 110, loss: 0.41029298305511475
step: 120, loss: 0.2933495342731476
step: 130, loss: 0.31362685561180115
step: 140, loss: 0.3691026568412781
step: 150, loss: 0.31595346331596375
step: 160, loss: 0.3848772943019867
step: 170, loss: 0.32622745633125305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2013793103448276, f1=0.2112482853223594,
best_f1=0.2112482853223594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31469863653182983
step: 10, loss: 0.42508962750434875
step: 20, loss: 0.31328901648521423
step: 30, loss: 0.31280753016471863
step: 40, loss: 0.06802205741405487
step: 50, loss: 0.3919774293899536
step: 60, loss: 0.14723755419254303
step: 70, loss: 0.4828985035419464
step: 80, loss: 0.257831871509552
step: 90, loss: 0.2630045711994171
step: 100, loss: 0.5559576153755188
step: 110, loss: 0.2849366068840027
step: 120, loss: 0.24444450438022614
step: 130, loss: 0.5729940533638
step: 140, loss: 0.5429772734642029
step: 150, loss: 0.4138830304145813
step: 160, loss: 0.3723691701889038
step: 170, loss: 0.26428326964378357
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6343612334801763, f1=0.6726057906458798,
best_f1=0.6726057906458798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3984614908695221
step: 10, loss: 0.14093108475208282
step: 20, loss: 0.22833652794361115
step: 30, loss: 0.17776019871234894
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.2876002788543701
step: 50, loss: 0.3835621178150177
step: 60, loss: 0.12165682017803192
step: 70, loss: 0.25888729095458984
step: 80, loss: 0.34995579719543457
step: 90, loss: 0.2968415319919586

```

```

step: 100, loss: 0.13162583112716675
step: 110, loss: 0.15286798775196075
step: 120, loss: 0.25439679622650146
step: 130, loss: 0.4690028131008148
step: 140, loss: 0.1927971988916397
step: 150, loss: 0.03743838146328926
step: 160, loss: 0.14649982750415802
step: 170, loss: 0.14532583951950073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7980295566502462, f1=0.7950617283950617,
best_f1=0.7950617283950617
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2787313461303711
step: 10, loss: 0.12059950828552246
step: 20, loss: 0.030409710481762886
step: 30, loss: 0.29521897435188293
step: 40, loss: 0.10029678046703339
step: 50, loss: 0.1668984740972519
step: 60, loss: 0.1676482856273651
step: 70, loss: 0.01611493155360222
step: 80, loss: 0.25917449593544006
step: 90, loss: 0.34156885743141174
step: 100, loss: 0.3113570213317871
step: 110, loss: 0.2064177244901657
step: 120, loss: 0.30503135919570923
step: 130, loss: 0.07863554358482361
step: 140, loss: 0.04964140057563782
step: 150, loss: 0.12924152612686157
step: 160, loss: 0.002833903068676591
step: 170, loss: 0.046065714210271835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8393285371702638, f1=0.836027713625866,
best_f1=0.836027713625866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05647197365760803
step: 10, loss: 0.0423557311296463
step: 20, loss: 0.1603495329618454
step: 30, loss: 0.042604342103004456
step: 40, loss: 0.1388978809118271
step: 50, loss: 0.04574083164334297
step: 60, loss: 0.061705928295850754
step: 70, loss: 0.07072868198156357
step: 80, loss: 0.003865661332383752
step: 90, loss: 0.03982335701584816
step: 100, loss: 0.12518282234668732
step: 110, loss: 0.0893532931804657
step: 120, loss: 0.14703163504600525
step: 130, loss: 0.012225530110299587
step: 140, loss: 0.08729628473520279
step: 150, loss: 0.08196165412664413
step: 160, loss: 0.02143843099474907
step: 170, loss: 0.023204665631055832
/usr/local/lib/python3.7/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=0.8258706467661691, f1=0.8619047619047621,
best_f1=0.836027713625866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17218060791492462
step: 10, loss: 0.04390424117445946
step: 20, loss: 0.09693292528390884
step: 30, loss: 0.011525500565767288
step: 40, loss: 0.04194042086601257
step: 50, loss: 0.015519052743911743
step: 60, loss: 0.07473268359899521
step: 70, loss: 0.016872605308890343
step: 80, loss: 0.019740069285035133
step: 90, loss: 0.20971795916557312
step: 100, loss: 0.029413998126983643
step: 110, loss: 0.17315229773521423
step: 120, loss: 0.09224406629800797
step: 130, loss: 0.047569431364536285
step: 140, loss: 0.03628778085112572
step: 150, loss: 0.014775331132113934
step: 160, loss: 0.03797513246536255
step: 170, loss: 0.04128990322351456
/usr/local/lib/python3.7/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=0.8309859154929576, f1=0.8482142857142858,
best_f1=0.836027713625866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16828829050064087
step: 10, loss: 0.05581571161746979
step: 20, loss: 0.004202584270387888
step: 30, loss: 0.012410396710038185
step: 40, loss: 0.0012211413122713566
step: 50, loss: 0.0026019783690571785
step: 60, loss: 0.005959305912256241
step: 70, loss: 0.05108754709362984
step: 80, loss: 0.047773949801921844
step: 90, loss: 0.039547767490148544
step: 100, loss: 0.0024112334940582514
step: 110, loss: 0.05437060818076134
step: 120, loss: 0.012801355682313442
step: 130, loss: 0.004501761868596077
step: 140, loss: 0.048715196549892426
step: 150, loss: 0.14661456644535065
step: 160, loss: 0.009755386039614677
step: 170, loss: 0.019951771944761276
/usr/local/lib/python3.7/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=0.835820895522388, f1=0.8815165876777251,
best_f1=0.836027713625866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07049702852964401
step: 10, loss: 0.02316814474761486
step: 20, loss: 0.01243654079735279
step: 30, loss: 0.0014452990144491196
step: 40, loss: 0.014165271073579788
step: 50, loss: 0.00607859343290329
step: 60, loss: 0.0037978114560246468
step: 70, loss: 0.05381698161363602
step: 80, loss: 0.006265606731176376
step: 90, loss: 0.06417122483253479
step: 100, loss: 0.0006552253034897149
step: 110, loss: 0.13082174956798553
step: 120, loss: 0.04801073670387268
step: 130, loss: 0.006391494069248438
step: 140, loss: 0.03739571571350098
step: 150, loss: 0.019074982032179832

```

```

step: 160, loss: 0.002317767823114991
step: 170, loss: 0.11060834676027298
/usr/local/lib/python3.7/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=0.8341463414634146, f1=0.8836104513064134,
best_f1=0.836027713625866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008555748499929905
step: 10, loss: 0.10097349435091019
step: 20, loss: 0.0025140251964330673
step: 30, loss: 0.04675140976905823
step: 40, loss: 0.008479767479002476
step: 50, loss: 0.003672860562801361
step: 60, loss: 0.04758870601654053
step: 70, loss: 0.15745411813259125
step: 80, loss: 0.006256467662751675
step: 90, loss: 0.08123473823070526
step: 100, loss: 0.020706739276647568
step: 110, loss: 0.016030021011829376
step: 120, loss: 0.00402096938341856
step: 130, loss: 0.038201604038476944
step: 140, loss: 0.04521356523036957
step: 150, loss: 0.22558793425559998
step: 160, loss: 0.007720176130533218
step: 170, loss: 0.017161088064312935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8415584415584415, f1=0.8780487804878048,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014007000718265772
step: 10, loss: 0.03614136949181557
step: 20, loss: 0.16359949111938477
step: 30, loss: 0.044044379144907
step: 40, loss: 0.013658496551215649
step: 50, loss: 0.01826772466301918
step: 60, loss: 0.0016187356086447835
step: 70, loss: 0.029156174510717392

```

```

step: 80, loss: 0.03407660871744156
step: 90, loss: 0.001718522747978568
step: 100, loss: 0.011361378245055676
step: 110, loss: 0.01235471572726965
step: 120, loss: 0.003224251326173544
step: 130, loss: 0.00990036129951477
step: 140, loss: 0.007741172332316637
step: 150, loss: 0.17985352873802185
step: 160, loss: 0.0026390538550913334
step: 170, loss: 0.0602865032851696
/usr/local/lib/python3.7/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=0.8377723970944311, f1=0.8611764705882352,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06601164489984512
step: 10, loss: 0.026423495262861252
step: 20, loss: 0.0014057527296245098
step: 30, loss: 0.019967537373304367
step: 40, loss: 0.004399226047098637
step: 50, loss: 0.0010559674119576812
step: 60, loss: 0.010918019339442253
step: 70, loss: 0.023455066606402397
step: 80, loss: 0.020823413506150246
step: 90, loss: 0.011893256567418575
step: 100, loss: 0.012776847928762436
step: 110, loss: 0.009463564492762089
step: 120, loss: 0.019097184762358665
step: 130, loss: 0.006541093345731497
step: 140, loss: 0.01342750433832407
step: 150, loss: 0.00132285850122571
step: 160, loss: 0.006035616155713797
step: 170, loss: 0.03436484560370445
/usr/local/lib/python3.7/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=0.8390243902439024, f1=0.8604651162790699,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008289706893265247
step: 10, loss: 0.06972043216228485
step: 20, loss: 0.0009420828428119421
step: 30, loss: 0.03425060957670212

```

```

step: 40, loss: 0.007743536960333586
step: 50, loss: 0.004303945694118738
step: 60, loss: 0.015097652561962605
step: 70, loss: 0.0005065252771601081
step: 80, loss: 0.0004818330053240061
step: 90, loss: 0.014919431880116463
step: 100, loss: 0.002379258396103978
step: 110, loss: 0.002111070090904832
step: 120, loss: 0.011856552213430405
step: 130, loss: 0.02835560403764248
step: 140, loss: 0.0006809932528994977
step: 150, loss: 0.0005920942639932036
step: 160, loss: 0.006824948824942112
step: 170, loss: 0.16987642645835876
/usr/local/lib/python3.7/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=0.837092731829574, f1=0.8481927710843373,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046250849845819175
step: 10, loss: 0.000638412544503808
step: 20, loss: 0.0006314957281574607
step: 30, loss: 0.07583095878362656
step: 40, loss: 0.005602964665740728
step: 50, loss: 0.0038000880740582943
step: 60, loss: 0.0009023857419379056
step: 70, loss: 0.10177125781774521
step: 80, loss: 0.000856181257404387
step: 90, loss: 0.02063494548201561
step: 100, loss: 0.01023964025080204
step: 110, loss: 0.0023888382129371166
step: 120, loss: 0.013601280748844147
step: 130, loss: 0.0013094589812681079
step: 140, loss: 0.019993001595139503
step: 150, loss: 0.000495040905661881
step: 160, loss: 0.0004422295023687184
step: 170, loss: 0.00028040833421982825
/usr/local/lib/python3.7/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=0.8393782383419689, f1=0.8850855745721272,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.006015298422425985
step: 10, loss: 0.0013304693857207894
step: 20, loss: 0.001293121138587594
step: 30, loss: 0.006977317854762077
step: 40, loss: 0.001471267663873732
step: 50, loss: 0.009527448564767838
step: 60, loss: 0.005542058032006025
step: 70, loss: 0.013579151593148708
step: 80, loss: 0.002846095012500882
step: 90, loss: 0.0003370062622707337
step: 100, loss: 0.0010250442428514361
step: 110, loss: 0.009886428713798523
step: 120, loss: 0.0003569510590750724
step: 130, loss: 0.11274275183677673
step: 140, loss: 0.014352121390402317
step: 150, loss: 0.03213309496641159
step: 160, loss: 0.007381456904113293
step: 170, loss: 0.03823774680495262
/usr/local/lib/python3.7/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=0.8374384236453202, f1=0.8699763593380615,
best_f1=0.8780487804878048
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008478416129946709
step: 10, loss: 0.0013112961314618587
step: 20, loss: 0.0004951838636770844
step: 30, loss: 0.0009026407497003675
step: 40, loss: 0.0003355748485773802
step: 50, loss: 0.00022143228852655739
step: 60, loss: 0.0033762860111892223
step: 70, loss: 0.11980410665273666
step: 80, loss: 0.0022264081053435802
step: 90, loss: 0.004983840510249138
step: 100, loss: 0.06684897094964981
step: 110, loss: 0.000525095674674958
step: 120, loss: 0.0003007413470186293
step: 130, loss: 0.000445683574071154
step: 140, loss: 0.013580644503235817
step: 150, loss: 0.00027493754168972373
step: 160, loss: 0.000671684683766216
step: 170, loss: 0.0003684184921439737
/usr/local/lib/python3.7/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=0.8374384236453202, f1=0.8720379146919431,

```

```

best_f1=0.8780487804878048
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:05, 322.76it/s]
load_f1 = 0.4535519125683059
real_f1 = 0.42446941323345816
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, 198.93it/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, 285kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 7.37MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 4.74MB/s]
Downloading: 100% 501M/501M [00:12<00:00, 40.3MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02

```



```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5656196475028992
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4660131633281708
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.4868883490562439
step: 30, loss: 0.08366415649652481
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.12873335182666779
step: 50, loss: 0.08870390057563782
step: 60, loss: 0.15612834692001343
step: 70, loss: 0.04164106026291847
step: 80, loss: 0.06330747902393341
step: 90, loss: 0.1253170669078827
step: 100, loss: 0.1858264058828354
step: 110, loss: 0.04389401525259018
step: 120, loss: 0.03243694454431534
step: 130, loss: 0.1359434276819229
step: 140, loss: 0.010591111145913601
step: 150, loss: 0.15822023153305054
step: 160, loss: 0.020832594484090805
step: 170, loss: 0.06338761001825333
step: 180, loss: 0.07245144993066788
step: 190, loss: 0.04624158516526222
step: 200, loss: 0.027390310540795326
step: 210, loss: 0.029834384098649025
step: 220, loss: 0.08317869901657104
step: 230, loss: 0.0021588450763374567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9534368070953436, f1=0.967741935483871,
best_f1=0.967741935483871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032420698553323746
step: 10, loss: 0.10824967920780182
step: 20, loss: 0.026509582996368408
step: 30, loss: 0.023895123973488808
step: 40, loss: 0.03534923493862152
step: 50, loss: 0.019439391791820526
step: 60, loss: 0.0024506051559001207
step: 70, loss: 0.017391296103596687
step: 80, loss: 0.010968505404889584
step: 90, loss: 0.0785694271326065
step: 100, loss: 0.04959284886717796
step: 110, loss: 0.07067383080720901
step: 120, loss: 0.007326112594455481
step: 130, loss: 0.013315496034920216
step: 140, loss: 0.002375976648181677
step: 150, loss: 0.01853661797940731
step: 160, loss: 0.03872659429907799
step: 170, loss: 0.0024661514908075333
step: 180, loss: 0.008509361185133457
step: 190, loss: 0.01260815653949976
step: 200, loss: 0.014788080006837845
step: 210, loss: 0.014349707402288914
step: 220, loss: 0.015847880393266678
step: 230, loss: 0.0017934620846062899
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9671574178935448, f1=0.9740112994350283,
best_f1=0.9740112994350283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04134608805179596
step: 10, loss: 0.017731575295329094
step: 20, loss: 0.002290265867486596
step: 30, loss: 0.0019027860835194588
step: 40, loss: 0.06365203112363815
step: 50, loss: 0.03183163329958916
step: 60, loss: 0.033488720655441284
step: 70, loss: 0.003511450020596385
step: 80, loss: 0.0012429393827915192
step: 90, loss: 0.032433539628982544

```

```

step: 100, loss: 0.03465230390429497
step: 110, loss: 0.015428347513079643
step: 120, loss: 0.0013652854831889272
step: 130, loss: 0.0077891782857477665
step: 140, loss: 0.002059027086943388
step: 150, loss: 0.009927196428179741
step: 160, loss: 0.002922221552580595
step: 170, loss: 0.0016707205213606358
step: 180, loss: 0.00692775659263134
step: 190, loss: 0.017316417768597603
step: 200, loss: 0.003554890863597393
step: 210, loss: 0.00227048946544528
step: 220, loss: 0.018296154215931892
step: 230, loss: 0.009331237524747849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9773755656108598, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005499598104506731
step: 10, loss: 0.0008232610416598618
step: 20, loss: 0.0038910910952836275
step: 30, loss: 0.001947490032762289
step: 40, loss: 0.06634963303804398
step: 50, loss: 0.004314634948968887
step: 60, loss: 0.0019045614171773195
step: 70, loss: 0.01489105075597763
step: 80, loss: 0.014499392360448837
step: 90, loss: 0.003191567724570632
step: 100, loss: 0.033718109130859375
step: 110, loss: 0.0009241968509741127
step: 120, loss: 0.00807754322886467
step: 130, loss: 0.002755837980657816
step: 140, loss: 0.0004581285174936056
step: 150, loss: 0.00038356747245416045
step: 160, loss: 0.001046070596203208
step: 170, loss: 0.0010164608247578144
step: 180, loss: 0.038701094686985016
step: 190, loss: 0.0008874474442563951
step: 200, loss: 0.013669311068952084
step: 210, loss: 0.050751570612192154

```

```

step: 220, loss: 0.019663147628307343
step: 230, loss: 0.011505963280797005
/usr/local/lib/python3.7/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=0.9754464285714286, f1=0.9842696629213483,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017771131824702024
step: 10, loss: 0.00472632609307766
step: 20, loss: 0.02000340446829796
step: 30, loss: 0.0006185855017974973
step: 40, loss: 0.008239123038947582
step: 50, loss: 0.0010116443736478686
step: 60, loss: 0.003916301764547825
step: 70, loss: 0.016414985060691833
step: 80, loss: 0.06822963804006577
step: 90, loss: 0.08192756026983261
step: 100, loss: 0.0003054745029658079
step: 110, loss: 0.029744751751422882
step: 120, loss: 0.014167220331728458
step: 130, loss: 0.0006793795619159937
step: 140, loss: 0.0012622536160051823
step: 150, loss: 0.04922455921769142
step: 160, loss: 0.0004098824574612081
step: 170, loss: 0.002687361091375351
step: 180, loss: 0.004359066020697355
step: 190, loss: 0.04392748326063156
step: 200, loss: 0.0064885979518294334
step: 210, loss: 0.008570393547415733
step: 220, loss: 0.0006412083166651428
step: 230, loss: 0.0008956685196608305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9788182831661093, f1=0.9797752808988766,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003601344069465995
step: 10, loss: 0.004924630280584097

```

```

step: 20, loss: 0.0009439336718060076
step: 30, loss: 0.0005129977944307029
step: 40, loss: 0.0002543505688663572
step: 50, loss: 0.0004767322971019894
step: 60, loss: 0.0015404688892886043
step: 70, loss: 0.16526161134243011
step: 80, loss: 0.0007662656134925783
step: 90, loss: 0.015149526298046112
step: 100, loss: 0.0012250124709680676
step: 110, loss: 0.01789373904466629
step: 120, loss: 0.0011277629528194666
step: 130, loss: 0.0013241431443020701
step: 140, loss: 0.0040932646952569485
step: 150, loss: 0.001124378526583314
step: 160, loss: 0.001359070185571909
step: 170, loss: 0.0010979330400004983
step: 180, loss: 0.0012052608653903008
step: 190, loss: 0.0010042025242000818
step: 200, loss: 0.002582474146038294
step: 210, loss: 0.0034822961315512657
step: 220, loss: 0.0018229454290121794
step: 230, loss: 0.0017229135846719146
/usr/local/lib/python3.7/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=0.9764309764309763, f1=0.9764837625979844,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004084428772330284
step: 10, loss: 0.0005033958586864173
step: 20, loss: 0.03625904768705368
step: 30, loss: 0.0006829790654592216
step: 40, loss: 0.0011327750980854034
step: 50, loss: 0.0023057060316205025
step: 60, loss: 0.0004916325560770929
step: 70, loss: 0.0006515830755233765
step: 80, loss: 0.0004202101845294237
step: 90, loss: 0.000391310837585479
step: 100, loss: 0.002734351670369506
step: 110, loss: 0.001346635865047574
step: 120, loss: 0.007178904954344034
step: 130, loss: 0.0064980811439454556
step: 140, loss: 0.00019807864737231284
step: 150, loss: 0.005880330689251423
step: 160, loss: 0.0002287999086547643
step: 170, loss: 0.025753073394298553

```

```

step: 180, loss: 0.00016464196960441768
step: 190, loss: 0.0017221367452293634
step: 200, loss: 0.00026439051725901663
step: 210, loss: 0.006880045402795076
step: 220, loss: 0.0002741737407632172
step: 230, loss: 0.00028259065584279597
/usr/local/lib/python3.7/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=0.9775280898876404,
best_f1=0.9797752808988766
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003450007061474025
step: 10, loss: 0.0010176385985687375
step: 20, loss: 0.008289402350783348
step: 30, loss: 0.0006395874079316854
step: 40, loss: 0.00089760881382972
step: 50, loss: 0.0009054560214281082
step: 60, loss: 0.0004194234497845173
step: 70, loss: 0.0004239371046423912
step: 80, loss: 0.01989850215613842
step: 90, loss: 0.0002853852929547429
step: 100, loss: 0.00023218843853101134
step: 110, loss: 0.0024232862051576376
step: 120, loss: 0.0004339100851211697
step: 130, loss: 0.00042668782407417893
step: 140, loss: 0.0009690944571048021
step: 150, loss: 0.17915542423725128
step: 160, loss: 0.017799338325858116
step: 170, loss: 0.13399043679237366
step: 180, loss: 0.001405645627528429
step: 190, loss: 0.0012498779688030481
step: 200, loss: 0.02620561420917511
step: 210, loss: 0.0005754899466410279
step: 220, loss: 0.0005725458613596857
step: 230, loss: 0.0002478020032867789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.979591836734694, 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: 0.00024956496781669557
step: 10, loss: 0.0001748797221807763
step: 20, loss: 0.0002364751126151532
step: 30, loss: 0.00032233790261670947
step: 40, loss: 0.00028470371034927666
step: 50, loss: 0.0003735333157237619
step: 60, loss: 0.00012580316979438066
step: 70, loss: 0.02283342555165291
step: 80, loss: 8.532965875929222e-05
step: 90, loss: 0.030928030610084534
step: 100, loss: 0.00010429038957227021
step: 110, loss: 0.00018172780983150005
step: 120, loss: 0.014523371122777462
step: 130, loss: 0.01631847769021988
step: 140, loss: 0.0001997720100916922
step: 150, loss: 0.004379601683467627
step: 160, loss: 0.004123387858271599
step: 170, loss: 0.0013837150763720274
step: 180, loss: 0.0005212260875850916
step: 190, loss: 0.00011650809028651565
step: 200, loss: 0.00044804144999943674
step: 210, loss: 0.0003182572836522013
step: 220, loss: 8.100292325252667e-05
step: 230, loss: 0.00010092508455272764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832026875699889, f1=0.9821029082774049,
best_f1=0.9821029082774049
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.738598262425512e-05
step: 10, loss: 0.00017196827684529126
step: 20, loss: 0.00018969207303598523
step: 30, loss: 5.584225073107518e-05
step: 40, loss: 0.0002470540057402104
step: 50, loss: 0.0006416678079403937
step: 60, loss: 0.00015451792569365352
step: 70, loss: 0.001123329158872366
step: 80, loss: 0.0002850664604920894
step: 90, loss: 0.0001521567755844444

```

```

step: 100, loss: 0.00014203341561369598
step: 110, loss: 0.004256194923073053
step: 120, loss: 7.118725625332445e-05
step: 130, loss: 9.498870349489152e-05
step: 140, loss: 6.884541653562337e-05
step: 150, loss: 0.0009473627433180809
step: 160, loss: 2.9156180971767753e-05
step: 170, loss: 6.405952444765717e-05
step: 180, loss: 0.0002692936977837235
step: 190, loss: 0.0001032094587571919
step: 200, loss: 0.00018243594968225807
step: 210, loss: 0.0387481153011322
step: 220, loss: 0.00018452864605933428
step: 230, loss: 9.091709944186732e-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=0.9865771812080537, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.255537351127714e-05
step: 10, loss: 0.00013103746459819376
step: 20, loss: 0.00019504186639096588
step: 30, loss: 7.753391400910914e-05
step: 40, loss: 2.1869238480576314e-05
step: 50, loss: 3.688437573146075e-05
step: 60, loss: 0.0012028934434056282
step: 70, loss: 6.215118628460914e-05
step: 80, loss: 8.40151114971377e-05
step: 90, loss: 0.02315818890929222
step: 100, loss: 0.0001287415507249534
step: 110, loss: 0.00018321500101592392
step: 120, loss: 7.425486546708271e-05
step: 130, loss: 0.0001209691763506271
step: 140, loss: 0.007682261988520622
step: 150, loss: 0.00028879891033284366
step: 160, loss: 0.05010821297764778
step: 170, loss: 0.0021081306040287018
step: 180, loss: 0.0006885818438604474
step: 190, loss: 0.0001745657209539786
step: 200, loss: 0.001259820768609643
step: 210, loss: 0.00017430844309274107

```



```

step: 220, loss: 0.0021852448116987944
step: 230, loss: 0.00016904593212530017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9865470852017937, f1=0.9753914988814317,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005079871043562889
step: 10, loss: 8.916010847315192e-05
step: 20, loss: 0.021971886977553368
step: 30, loss: 0.03809206187725067
step: 40, loss: 0.0002716888557188213
step: 50, loss: 0.0002816616033669561
step: 60, loss: 0.00013357926218304783
step: 70, loss: 7.494095916626975e-05
step: 80, loss: 6.991417467361316e-05
step: 90, loss: 0.004330918192863464
step: 100, loss: 5.575533214141615e-05
step: 110, loss: 0.0012444857275113463
step: 120, loss: 9.639470226829872e-05
step: 130, loss: 0.00011839559738291427
step: 140, loss: 0.00028830018709413707
step: 150, loss: 6.390845373971388e-05
step: 160, loss: 0.00011666679347399622
step: 170, loss: 0.0002048325550276786
step: 180, loss: 0.00033945750328712165
step: 190, loss: 0.003415230894461274
step: 200, loss: 0.00013221967674326152
step: 210, loss: 0.0027531476225703955
step: 220, loss: 0.006095534190535545
step: 230, loss: 0.00022584726684726775
/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9808342728297633,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000733793480321765
step: 10, loss: 0.0005778488703072071
step: 20, loss: 9.14669144549407e-05
step: 30, loss: 3.9624654164072126e-05
step: 40, loss: 0.00016227959713432938
step: 50, loss: 0.00037307688035070896

```

```

step: 60, loss: 9.495475387666374e-05
step: 70, loss: 6.938860315131024e-05
step: 80, loss: 4.82782197650522e-05
step: 90, loss: 6.127360393293202e-05
step: 100, loss: 0.00021214182197581977
step: 110, loss: 0.0006588619435206056
step: 120, loss: 5.3061481594340876e-05
step: 130, loss: 0.00010954665776807815
step: 140, loss: 7.826091314200312e-05
step: 150, loss: 2.8805568945244886e-05
step: 160, loss: 0.018148163333535194
step: 170, loss: 0.00010907107207458466
step: 180, loss: 0.05049543455243111
step: 190, loss: 7.81866256147623e-05
step: 200, loss: 7.875248411437497e-05
step: 210, loss: 0.0005712545244023204
step: 220, loss: 0.00010925389506155625
step: 230, loss: 4.401515980134718e-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=0.9844097995545658, f1=0.9776785714285714,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.7619473005179316e-05
step: 10, loss: 4.0427759813610464e-05
step: 20, loss: 0.00014742073835805058
step: 30, loss: 7.444398943334818e-05
step: 40, loss: 6.50522779324092e-05
step: 50, loss: 3.485918932710774e-05
step: 60, loss: 4.737931885756552e-05
step: 70, loss: 3.135437145829201e-05
step: 80, loss: 5.0326827476965263e-05
step: 90, loss: 7.347347855102271e-05
step: 100, loss: 4.7406185331055894e-05
step: 110, loss: 4.2935440433211625e-05
step: 120, loss: 1.9538745618774556e-05
step: 130, loss: 7.502375228796154e-05
step: 140, loss: 0.0006260992959141731
step: 150, loss: 6.145537918200716e-05
step: 160, loss: 4.836415610043332e-05
step: 170, loss: 3.867386112688109e-05
step: 180, loss: 3.702761750901118e-05
step: 190, loss: 2.9771350455121137e-05
step: 200, loss: 5.4622982133878395e-05
step: 210, loss: 3.16708319587633e-05

```

```

step: 220, loss: 4.16831680922769e-05
step: 230, loss: 2.452283115417231e-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=0.9833147942157954, f1=0.9777777777777777,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022942088544368744
step: 10, loss: 4.039453051518649e-05
step: 20, loss: 0.00017857116472441703
step: 30, loss: 4.938496567774564e-05
step: 40, loss: 2.5606927010812797e-05
step: 50, loss: 3.381635178811848e-05
step: 60, loss: 0.03532925993204117
step: 70, loss: 4.380282916827127e-05
step: 80, loss: 3.3805576094891876e-05
step: 90, loss: 2.9901599191362038e-05
step: 100, loss: 2.4880275304894894e-05
step: 110, loss: 3.9799069782020524e-05
step: 120, loss: 0.02210344187915325
step: 130, loss: 2.590475560282357e-05
step: 140, loss: 0.011825377121567726
step: 150, loss: 0.00010890424164244905
step: 160, loss: 0.010859152302145958
step: 170, loss: 2.0346940800664015e-05
step: 180, loss: 4.6905362978577614e-05
step: 190, loss: 4.027930481242947e-05
step: 200, loss: 0.0005455499049276114
step: 210, loss: 7.369189552264288e-05
step: 220, loss: 3.979100802098401e-05
step: 230, loss: 5.223970947554335e-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=0.9833147942157954, f1=0.9778270509977827,
best_f1=0.9788182831661093
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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, 186.70it/s]
load_f1 = 0.983277591973244
real_f1 = 0.9821826280623607
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:25, 172.60it/s]

```

1.3.2 DBLP-GoogleScholar - Running the matcher

```

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

```

```

--da attr_del \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

```
step: 0, loss: 0.6222003698348999
```

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: 0.441715806722641
```

```
step: 20, loss: 0.4924459457397461
```

step: 30, loss: 0.33304744958877563
step: 40, loss: 0.41994935274124146
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.7814359068870544
step: 60, loss: 0.35123974084854126
step: 70, loss: 0.4355365037918091
step: 80, loss: 0.46261829137802124
step: 90, loss: 0.2571801245212555
step: 100, loss: 0.6627786159515381
step: 110, loss: 0.2367372065782547
step: 120, loss: 0.33518582582473755
step: 130, loss: 0.1527250111103058
step: 140, loss: 0.2353198230266571
step: 150, loss: 0.12345597892999649
step: 160, loss: 0.26836690306663513
step: 170, loss: 0.2567014992237091
step: 180, loss: 0.196523979306221
step: 190, loss: 0.18758907914161682
step: 200, loss: 0.061625104397535324
step: 210, loss: 0.07351146638393402
step: 220, loss: 0.07135337591171265
step: 230, loss: 0.5750873684883118
step: 240, loss: 0.03048415668308735
step: 250, loss: 0.07524757832288742
step: 260, loss: 0.1400759369134903
step: 270, loss: 0.331915944814682
step: 280, loss: 0.07177291810512543
step: 290, loss: 0.14740031957626343
step: 300, loss: 0.16422775387763977
step: 310, loss: 0.17941312491893768
step: 320, loss: 0.20255866646766663
step: 330, loss: 0.12120068073272705
step: 340, loss: 0.48600852489471436
step: 350, loss: 0.09685970842838287
step: 360, loss: 0.04459739476442337
step: 370, loss: 0.13143205642700195
step: 380, loss: 0.1572815626859665
step: 390, loss: 0.028033792972564697
step: 400, loss: 0.05699659138917923
step: 410, loss: 0.23722288012504578
step: 420, loss: 0.006569257006049156
step: 430, loss: 0.04479324072599411
step: 440, loss: 0.04952926188707352
step: 450, loss: 0.0214739590883255
step: 460, loss: 0.03205506131052971
step: 470, loss: 0.03296695277094841
step: 480, loss: 0.15669885277748108
step: 490, loss: 0.11744147539138794

```

step: 500, loss: 0.030296484008431435
step: 510, loss: 0.04555335268378258
step: 520, loss: 0.19451528787612915
step: 530, loss: 0.03641543537378311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.936500685244404, f1=0.9372997711670481,
best_f1=0.9372997711670481
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07717806845903397
step: 10, loss: 0.06106080487370491
step: 20, loss: 0.028894662857055664
step: 30, loss: 0.05577109009027481
step: 40, loss: 0.038158293813467026
step: 50, loss: 0.056159403175115585
step: 60, loss: 0.021499717608094215
step: 70, loss: 0.0420832522213459
step: 80, loss: 0.0346209779381752
step: 90, loss: 0.0034896230790764093
step: 100, loss: 0.09689649939537048
step: 110, loss: 0.03106559067964554
step: 120, loss: 0.06223531812429428
step: 130, loss: 0.010248127393424511
step: 140, loss: 0.10033629834651947
step: 150, loss: 0.034442588686943054
step: 160, loss: 0.051040057092905045
step: 170, loss: 0.02826656401157379
step: 180, loss: 0.0550132617354393
step: 190, loss: 0.015152274630963802
step: 200, loss: 0.11218488961458206
step: 210, loss: 0.01226609107106924
step: 220, loss: 0.00044327080831862986
step: 230, loss: 0.049217768013477325
step: 240, loss: 0.04408353567123413
step: 250, loss: 0.010847114026546478
step: 260, loss: 0.09533525258302689
step: 270, loss: 0.007554061245173216
step: 280, loss: 0.05889861658215523
step: 290, loss: 0.10224807262420654
step: 300, loss: 0.07469760626554489
step: 310, loss: 0.07006050646305084

```

```

step: 320, loss: 0.04072677344083786
step: 330, loss: 0.06719998270273209
step: 340, loss: 0.08976215869188309
step: 350, loss: 0.0018049292266368866
step: 360, loss: 0.046431466937065125
step: 370, loss: 0.009305363520979881
step: 380, loss: 0.15228481590747833
step: 390, loss: 0.017586465924978256
step: 400, loss: 0.06755095720291138
step: 410, loss: 0.03580177202820778
step: 420, loss: 0.10589861124753952
step: 430, loss: 0.2509503662586212
step: 440, loss: 0.04195942357182503
step: 450, loss: 0.03861679136753082
step: 460, loss: 0.07330061495304108
step: 470, loss: 0.058983515948057175
step: 480, loss: 0.16416160762310028
step: 490, loss: 0.06655442714691162
step: 500, loss: 0.0011734466534107924
step: 510, loss: 0.01857457496225834
step: 520, loss: 0.4513237178325653
step: 530, loss: 0.09429498761892319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488847583643123, f1=0.9506517690875234,
best_f1=0.9506517690875234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11710542440414429
step: 10, loss: 0.25201526284217834
step: 20, loss: 0.038692500442266464
step: 30, loss: 0.060469288378953934
step: 40, loss: 0.10679803043603897
step: 50, loss: 0.008788339793682098
step: 60, loss: 0.01812152937054634
step: 70, loss: 0.023801032453775406
step: 80, loss: 0.006028974894434214
step: 90, loss: 0.0456664077937603
step: 100, loss: 0.06848964840173721
step: 110, loss: 0.029294367879629135
step: 120, loss: 0.08392760157585144
step: 130, loss: 0.057519957423210144

```



```

step: 140, loss: 0.0323600098490715
step: 150, loss: 0.059839408844709396
step: 160, loss: 0.02729255147278309
step: 170, loss: 0.005191660486161709
step: 180, loss: 0.04036259651184082
step: 190, loss: 0.006797819398343563
step: 200, loss: 0.15322180092334747
step: 210, loss: 0.036711618304252625
step: 220, loss: 0.08723486214876175
step: 230, loss: 0.030726274475455284
step: 240, loss: 0.04221133142709732
step: 250, loss: 0.21571536362171173
step: 260, loss: 0.13272713124752045
step: 270, loss: 0.008177216164767742
step: 280, loss: 0.005370055790990591
step: 290, loss: 0.010808294638991356
step: 300, loss: 0.1764201521873474
step: 310, loss: 0.023436589166522026
step: 320, loss: 0.013806188479065895
step: 330, loss: 0.006181950215250254
step: 340, loss: 0.023930134251713753
step: 350, loss: 0.1095363050699234
step: 360, loss: 0.016216682270169258
step: 370, loss: 0.12002037465572357
step: 380, loss: 0.02451488748192787
step: 390, loss: 0.019559752196073532
step: 400, loss: 0.1440357118844986
step: 410, loss: 0.021690459921956062
step: 420, loss: 0.010823423974215984
step: 430, loss: 0.03001580946147442
step: 440, loss: 0.26056990027427673
step: 450, loss: 0.07195199280977249
step: 460, loss: 0.10958007723093033
step: 470, loss: 0.07198704034090042
step: 480, loss: 0.18790769577026367
step: 490, loss: 0.05849718675017357
step: 500, loss: 0.03286411985754967
step: 510, loss: 0.008434086106717587
step: 520, loss: 0.025561166927218437
step: 530, loss: 0.015844060108065605
/usr/local/lib/python3.7/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=0.9480401093892434, f1=0.9419295839048926,
best_f1=0.9506517690875234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.04739604517817497
step: 10, loss: 0.043449901044368744
step: 20, loss: 0.04030688852071762
step: 30, loss: 0.1119546890258789
step: 40, loss: 0.03500993177294731
step: 50, loss: 0.05199700966477394
step: 60, loss: 0.011940685100853443
step: 70, loss: 0.022839175537228584
step: 80, loss: 0.009028027765452862
step: 90, loss: 0.1354503184556961
step: 100, loss: 0.006035300903022289
step: 110, loss: 0.09876503050327301
step: 120, loss: 0.0007596822688356042
step: 130, loss: 0.14476114511489868
step: 140, loss: 0.011297729797661304
step: 150, loss: 0.0070063164457678795
step: 160, loss: 0.01185784861445427
step: 170, loss: 0.015259088948369026
step: 180, loss: 0.09460444003343582
step: 190, loss: 0.06151518598198891
step: 200, loss: 0.0450700968503952
step: 210, loss: 0.0026203494053333998
step: 220, loss: 0.010992485098540783
step: 230, loss: 0.009639492258429527
step: 240, loss: 0.037663962692022324
step: 250, loss: 0.20740796625614166
step: 260, loss: 0.007315068040043116
step: 270, loss: 0.07504472881555557
step: 280, loss: 0.00274072028696537
step: 290, loss: 0.009875461459159851
step: 300, loss: 0.0057718511670827866
step: 310, loss: 0.028743546456098557
step: 320, loss: 0.021006990224123
step: 330, loss: 0.2133132368326187
step: 340, loss: 0.08073447644710541
step: 350, loss: 0.06883871555328369
step: 360, loss: 0.05073815956711769
step: 370, loss: 0.08870874345302582
step: 380, loss: 0.0029117949306964874
step: 390, loss: 0.000447123748017475
step: 400, loss: 0.01419015135616064
step: 410, loss: 0.002469543367624283
step: 420, loss: 0.033103566616773605
step: 430, loss: 0.040252141654491425
step: 440, loss: 0.008610755205154419
step: 450, loss: 0.09476059675216675
step: 460, loss: 0.1585431545972824
step: 470, loss: 0.003992822021245956

```

step: 480, loss: 0.021976754069328308
step: 490, loss: 0.002023632638156414
step: 500, loss: 0.01041758619248867
step: 510, loss: 0.02385885827243328
step: 520, loss: 0.06828002631664276
step: 530, loss: 0.10824213176965714
/usr/local/lib/python3.7/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=0.9463459759481962, f1=0.9419953596287702,
best_f1=0.9506517690875234
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07753412425518036
step: 10, loss: 0.05717173591256142
step: 20, loss: 0.01795085333287716
step: 30, loss: 0.006890794727951288
step: 40, loss: 0.0028900946490466595
step: 50, loss: 0.07848761230707169
step: 60, loss: 0.04101048782467842
step: 70, loss: 0.009379367344081402
step: 80, loss: 0.005353809799998999
step: 90, loss: 0.10151024162769318
step: 100, loss: 0.04491940140724182
step: 110, loss: 0.005721325520426035
step: 120, loss: 0.1101977601647377
step: 130, loss: 0.020509082823991776
step: 140, loss: 0.006080156657844782
step: 150, loss: 0.016865288838744164
step: 160, loss: 0.045570533722639084
step: 170, loss: 0.13874687254428864
step: 180, loss: 0.07126989960670471
step: 190, loss: 0.007663455791771412
step: 200, loss: 0.028912357985973358
step: 210, loss: 0.0022939748596400023
step: 220, loss: 0.01287966500967741
step: 230, loss: 0.005510437302291393
step: 240, loss: 0.0022630884777754545
step: 250, loss: 0.1223568543791771
step: 260, loss: 0.006601425819098949
step: 270, loss: 0.009558686055243015
step: 280, loss: 0.002546766772866249
step: 290, loss: 0.009796754457056522
step: 300, loss: 0.009708092547953129
step: 310, loss: 0.05263601988554001
step: 320, loss: 0.05337977781891823
step: 330, loss: 0.0013199448585510254

```

```

step: 340, loss: 0.005081729032099247
step: 350, loss: 0.0014005015837028623
step: 360, loss: 0.0003784721193369478
step: 370, loss: 0.02423756942152977
step: 380, loss: 0.0006305929855443537
step: 390, loss: 0.012183014303445816
step: 400, loss: 0.02864570915699005
step: 410, loss: 0.05868351832032204
step: 420, loss: 0.20795665681362152
step: 430, loss: 0.04806753247976303
step: 440, loss: 0.009312148205935955
step: 450, loss: 0.004981997888535261
step: 460, loss: 0.03523017838597298
step: 470, loss: 0.01852680929005146
step: 480, loss: 0.035366788506507874
step: 490, loss: 0.0538986474275589
step: 500, loss: 0.005832454655319452
step: 510, loss: 0.16078144311904907
step: 520, loss: 0.09488686919212341
step: 530, loss: 0.07075633853673935
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9501385041551247, f1=0.950622981079834,
best_f1=0.950622981079834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05965462327003479
step: 10, loss: 0.006667339708656073
step: 20, loss: 0.005952776875346899
step: 30, loss: 0.003782388288527727
step: 40, loss: 0.0025530692655593157
step: 50, loss: 0.0026953844353556633
step: 60, loss: 0.03069491684436798
step: 70, loss: 0.05425788834691048
step: 80, loss: 8.761856588535011e-05
step: 90, loss: 0.003128078067675233
step: 100, loss: 0.002871121745556593
step: 110, loss: 0.000705498328898102
step: 120, loss: 0.03675683215260506
step: 130, loss: 0.007054369430989027
step: 140, loss: 0.0031176996417343616
step: 150, loss: 0.0028118775226175785

```

```
step: 160, loss: 0.10564939677715302
step: 170, loss: 0.025153808295726776
step: 180, loss: 0.06610502302646637
step: 190, loss: 0.06806498765945435
step: 200, loss: 0.004040305502712727
step: 210, loss: 0.007363086100667715
step: 220, loss: 0.002204928547143936
step: 230, loss: 0.010421528480947018
step: 240, loss: 0.008393976837396622
step: 250, loss: 0.019223148003220558
step: 260, loss: 0.011376067996025085
step: 270, loss: 0.027352668344974518
step: 280, loss: 0.005060754716396332
step: 290, loss: 0.007190273609012365
step: 300, loss: 0.005807032343000174
step: 310, loss: 0.03679579123854637
step: 320, loss: 0.0006044707261025906
step: 330, loss: 0.00725104846060276
step: 340, loss: 0.0012709763832390308
step: 350, loss: 0.010770034044981003
step: 360, loss: 0.07799990475177765
step: 370, loss: 0.004698641132563353
step: 380, loss: 0.0006968314992263913
step: 390, loss: 0.001962907612323761
step: 400, loss: 0.02129766158759594
step: 410, loss: 0.0011866531567648053
step: 420, loss: 0.002928289119154215
step: 430, loss: 0.0024120421148836613
step: 440, loss: 0.0011226021451875567
step: 450, loss: 0.17470240592956543
step: 460, loss: 0.0007306014886125922
step: 470, loss: 0.001416803919710219
step: 480, loss: 0.010909968987107277
step: 490, loss: 0.03343057259917259
step: 500, loss: 0.00775984488427639
step: 510, loss: 0.20192380249500275
step: 520, loss: 0.0002126903273165226
step: 530, loss: 0.002163185505196452
/usr/local/lib/python3.7/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=0.9450651769087525, f1=0.9381107491856677,
best_f1=0.950622981079834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03152158856391907
step: 10, loss: 0.0015051180962473154
```

step: 20, loss: 0.031090406700968742
step: 30, loss: 0.015242014080286026
step: 40, loss: 0.002448048908263445
step: 50, loss: 0.019067110493779182
step: 60, loss: 0.03986773639917374
step: 70, loss: 0.002766034798696637
step: 80, loss: 0.00859682634472847
step: 90, loss: 0.001805643318220973
step: 100, loss: 0.02611522376537323
step: 110, loss: 0.0001278874697163701
step: 120, loss: 0.005009929183870554
step: 130, loss: 0.0004611740296240896
step: 140, loss: 0.0027598626911640167
step: 150, loss: 0.001988770207390189
step: 160, loss: 0.0011032947804778814
step: 170, loss: 0.015080038458108902
step: 180, loss: 0.013807366602122784
step: 190, loss: 0.002238259185105562
step: 200, loss: 0.00022167968563735485
step: 210, loss: 0.009348790161311626
step: 220, loss: 0.00016891161794774234
step: 230, loss: 0.0004430313711054623
step: 240, loss: 0.0013800222659483552
step: 250, loss: 0.008702903985977173
step: 260, loss: 0.0027158271986991167
step: 270, loss: 0.005016086623072624
step: 280, loss: 0.15333129465579987
step: 290, loss: 0.009794702753424644
step: 300, loss: 0.0024055021349340677
step: 310, loss: 0.0026642398443073034
step: 320, loss: 0.2032759040594101
step: 330, loss: 0.013797281309962273
step: 340, loss: 0.0047060162760317326
step: 350, loss: 0.0013222777051851153
step: 360, loss: 0.005593487527221441
step: 370, loss: 0.007307325955480337
step: 380, loss: 0.006384361535310745
step: 390, loss: 0.01859114319086075
step: 400, loss: 0.042685288935899734
step: 410, loss: 0.0004268219927325845
step: 420, loss: 0.032825302332639694
step: 430, loss: 0.0005534613155759871
step: 440, loss: 0.020845046266913414
step: 450, loss: 0.00227339961566031
step: 460, loss: 0.00038632869836874306
step: 470, loss: 0.21244893968105316
step: 480, loss: 0.002954704686999321
step: 490, loss: 0.004926327150315046

```

step: 500, loss: 0.013646638952195644
step: 510, loss: 0.000816230836790055
step: 520, loss: 0.004010355100035667
step: 530, loss: 0.007686976809054613
/usr/local/lib/python3.7/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=0.9468779123951537, f1=0.9422180801491147,
best_f1=0.950622981079834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014819559874013066
step: 10, loss: 0.013073506765067577
step: 20, loss: 0.05241136625409126
step: 30, loss: 0.01282019354403019
step: 40, loss: 0.0013046835083514452
step: 50, loss: 0.002033766359090805
step: 60, loss: 0.005479909013956785
step: 70, loss: 0.007757318671792746
step: 80, loss: 0.01595298759639263
step: 90, loss: 0.00041648189653642476
step: 100, loss: 0.057923510670661926
step: 110, loss: 0.0009327879524789751
step: 120, loss: 0.007615454029291868
step: 130, loss: 0.021716393530368805
step: 140, loss: 0.0027786698192358017
step: 150, loss: 0.006821414455771446
step: 160, loss: 0.0027977062854915857
step: 170, loss: 0.057744454592466354
step: 180, loss: 0.1154170110821724
step: 190, loss: 0.10711006075143814
step: 200, loss: 0.0029255119152367115
step: 210, loss: 0.0254681259393692
step: 220, loss: 0.01988564245402813
step: 230, loss: 0.03027408942580223
step: 240, loss: 0.0524652861058712
step: 250, loss: 0.0007999019580893219
step: 260, loss: 0.0015832737553864717
step: 270, loss: 0.002658577635884285
step: 280, loss: 0.001125961891375482
step: 290, loss: 0.011709967628121376
step: 300, loss: 0.0025126615073531866
step: 310, loss: 0.007905748672783375
step: 320, loss: 0.0006185239762999117
step: 330, loss: 0.0005050793988630176
step: 340, loss: 0.08107347786426544
step: 350, loss: 0.0010072094155475497

```

```

step: 360, loss: 0.04360383376479149
step: 370, loss: 0.058767396956682205
step: 380, loss: 0.00039038588874973357
step: 390, loss: 0.0007648924947716296
step: 400, loss: 0.004822574555873871
step: 410, loss: 0.031193751841783524
step: 420, loss: 0.011666159145534039
step: 430, loss: 0.011228418909013271
step: 440, loss: 0.005488245747983456
step: 450, loss: 0.0008689977694302797
step: 460, loss: 0.014150053262710571
step: 470, loss: 0.004725296515971422
step: 480, loss: 0.0026447486598044634
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.01471002958714962
step: 500, loss: 0.0006870990619063377
step: 510, loss: 0.021878886967897415
step: 520, loss: 0.006095278542488813
step: 530, loss: 0.019209619611501694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9506976744186046, f1=0.9492787342950209,
best_f1=0.9492787342950209
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009204355883412063
step: 10, loss: 0.013479391112923622
step: 20, loss: 0.00010825283970916644
step: 30, loss: 0.13837674260139465
step: 40, loss: 0.0002073893410852179
step: 50, loss: 9.712594328448176e-05
step: 60, loss: 0.0012848003534600139
step: 70, loss: 0.10620344430208206
step: 80, loss: 0.003620308358222246
step: 90, loss: 0.0052806013263762
step: 100, loss: 0.003778391284868121
step: 110, loss: 0.06639818102121353
step: 120, loss: 0.0023482218384742737
step: 130, loss: 0.0046451822854578495
step: 140, loss: 0.002947506494820118
step: 150, loss: 0.009181030094623566
step: 160, loss: 0.0003875725087709725

```



```

step: 170, loss: 0.00027817863156087697
step: 180, loss: 0.0012449692003428936
step: 190, loss: 6.812810170231387e-05
step: 200, loss: 0.001125498441979289
step: 210, loss: 0.007778756320476532
step: 220, loss: 0.00046041555469855666
step: 230, loss: 0.0022186175920069218
step: 240, loss: 0.01972826197743416
step: 250, loss: 0.03082927130162716
step: 260, loss: 0.0002367007255088538
step: 270, loss: 0.006129749119281769
step: 280, loss: 0.015758275985717773
step: 290, loss: 0.0033875461667776108
step: 300, loss: 0.00019240194524172693
step: 310, loss: 0.01351465005427599
step: 320, loss: 0.03181695565581322
step: 330, loss: 0.0007771234377287328
step: 340, loss: 0.01061537116765976
step: 350, loss: 0.007268356624990702
step: 360, loss: 0.0008063826244324446
step: 370, loss: 0.0005256272270344198
step: 380, loss: 0.0003399711276870221
step: 390, loss: 0.015238259918987751
step: 400, loss: 0.02014285698533058
step: 410, loss: 0.0027413987554609776
step: 420, loss: 0.0013462848728522658
step: 430, loss: 0.016691964119672775
step: 440, loss: 0.0014688489027321339
step: 450, loss: 0.06481009721755981
step: 460, loss: 0.004591918550431728
step: 470, loss: 0.0004568478325381875
step: 480, loss: 0.0004229435289744288
step: 490, loss: 0.016878793016076088
step: 500, loss: 0.0035499532241374254
step: 510, loss: 0.001810066169127822
step: 520, loss: 0.010665002278983593
step: 530, loss: 0.0074084000661969185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9508949059201468, f1=0.9458218549127642,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019164301920682192
step: 10, loss: 0.001083008712157607
step: 20, loss: 0.0016792190726846457
step: 30, loss: 0.033430807292461395
step: 40, loss: 0.0007917886250652373
step: 50, loss: 0.001307420781813562
step: 60, loss: 0.0006265125120989978
step: 70, loss: 0.014394883066415787
step: 80, loss: 0.005029833875596523
step: 90, loss: 0.0032051268499344587
step: 100, loss: 0.028410611674189568
step: 110, loss: 0.009195182472467422
step: 120, loss: 0.00027334498008713126
step: 130, loss: 0.005797244608402252
step: 140, loss: 0.011556481942534447
step: 150, loss: 0.0018779640085995197
step: 160, loss: 0.07411372661590576
step: 170, loss: 0.0017446420388296247
step: 180, loss: 0.01289571076631546
step: 190, loss: 0.0007228812901303172
step: 200, loss: 0.003081511938944459
step: 210, loss: 0.0021652968134731054
step: 220, loss: 0.014618354849517345
step: 230, loss: 0.018966982141137123
step: 240, loss: 0.02032732591032982
step: 250, loss: 0.017736442387104034
step: 260, loss: 0.017471879720687866
step: 270, loss: 0.004879772663116455
step: 280, loss: 0.0064222682267427444
step: 290, loss: 4.1872837755363435e-05
step: 300, loss: 0.00030045799212530255
step: 310, loss: 0.07593141496181488
step: 320, loss: 0.05721817910671234
step: 330, loss: 0.003143774811178446
step: 340, loss: 0.00016253527428489178
step: 350, loss: 0.0015350835165008903
step: 360, loss: 7.247865141835064e-05
step: 370, loss: 0.00033840694231912494
step: 380, loss: 0.001100230379961431
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 390, loss: 0.0003749120223801583
step: 400, loss: 0.004190603271126747
step: 410, loss: 0.0018958377186208963
step: 420, loss: 0.001030571642331779
step: 430, loss: 0.0011655575362965465
step: 440, loss: 0.0008834078907966614
step: 450, loss: 0.04641331359744072
```

```

step: 460, loss: 0.0007037464529275894
step: 470, loss: 0.0069305626675486565
step: 480, loss: 0.0004943586536683142
step: 490, loss: 0.00018553207337390631
step: 500, loss: 0.0609208345413208
step: 510, loss: 0.09233269840478897
step: 520, loss: 0.012706940062344074
step: 530, loss: 0.04854907467961311
/usr/local/lib/python3.7/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=0.9491059147180192, f1=0.9483870967741935,
best_f1=0.9458218549127642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014122361317276955
step: 10, loss: 0.002339547500014305
step: 20, loss: 0.0007234174408949912
step: 30, loss: 0.0010938395280390978
step: 40, loss: 0.0001272177032660693
step: 50, loss: 0.0006710385205224156
step: 60, loss: 0.00304593937471509
step: 70, loss: 0.00025857577566057444
step: 80, loss: 0.00022106985852587968
step: 90, loss: 0.004227744415402412
step: 100, loss: 0.006919851526618004
step: 110, loss: 0.002576029859483242
step: 120, loss: 0.0015437697293236852
step: 130, loss: 0.0003054687986150384
step: 140, loss: 0.000933634175453335
step: 150, loss: 0.0027762376703321934
step: 160, loss: 0.0009344230638816953
step: 170, loss: 0.0032256797421723604
step: 180, loss: 0.0039882780984044075
step: 190, loss: 0.0010455651208758354
step: 200, loss: 0.00020397795015014708
step: 210, loss: 0.00037499156314879656
step: 220, loss: 0.09934256225824356
step: 230, loss: 0.00019497532048262656
step: 240, loss: 0.0008906862931326032
step: 250, loss: 0.004528927616775036
step: 260, loss: 0.0016332932282239199
step: 270, loss: 0.001677064341492951
step: 280, loss: 0.0006572139100171626
step: 290, loss: 0.0006268281722441316
step: 300, loss: 0.0001839076867327094
step: 310, loss: 0.001976663013920188

```

```

step: 320, loss: 0.0022795884869992733
step: 330, loss: 0.0001711669028736651
step: 340, loss: 0.0022351753432303667
step: 350, loss: 0.02959066443145275
step: 360, loss: 0.020421897992491722
step: 370, loss: 0.0024043829180300236
step: 380, loss: 0.002250832971185446
step: 390, loss: 0.001168757095001638
step: 400, loss: 0.001502186176367104
step: 410, loss: 0.0018492293311282992
step: 420, loss: 0.0013814553385600448
step: 430, loss: 0.0011554937809705734
step: 440, loss: 0.0028129485435783863
step: 450, loss: 0.011935155838727951
step: 460, loss: 0.029293503612279892
step: 470, loss: 0.06501088291406631
step: 480, loss: 0.007822005078196526
step: 490, loss: 0.00026470451848581433
step: 500, loss: 0.000284994428511709
step: 510, loss: 0.0015077688731253147
step: 520, loss: 0.004333769902586937
step: 530, loss: 0.015789011493325233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9530887134231305, f1=0.9480037140204272,
best_f1=0.9480037140204272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008939491235651076
step: 10, loss: 0.0018325322307646275
step: 20, loss: 0.0006958711892366409
step: 30, loss: 0.0006884181639179587
step: 40, loss: 0.09592515230178833
step: 50, loss: 0.027850134298205376
step: 60, loss: 0.002269444987177849
step: 70, loss: 0.0006261668750084937
step: 80, loss: 0.0001125319249695167
step: 90, loss: 0.03547954186797142
step: 100, loss: 0.014048442244529724
step: 110, loss: 0.00943106971681118
step: 120, loss: 9.571962436893955e-05
step: 130, loss: 0.0009125571814365685

```

```

step: 140, loss: 4.35315232607536e-05
step: 150, loss: 6.775380461476743e-05
step: 160, loss: 0.0020973384380340576
step: 170, loss: 5.727450843551196e-05
step: 180, loss: 0.0016961764777079225
step: 190, loss: 0.00023578290711157024
step: 200, loss: 0.0005145330214872956
step: 210, loss: 0.0004827285301871598
step: 220, loss: 3.452484816079959e-05
step: 230, loss: 0.005675602238625288
step: 240, loss: 0.0011103940196335316
step: 250, loss: 3.145938171655871e-05
step: 260, loss: 0.00794743001461029
step: 270, loss: 0.0008342845248989761
step: 280, loss: 0.00353570026345551
step: 290, loss: 0.0043984572403132915
step: 300, loss: 0.0005032566259615123
step: 310, loss: 0.0005198511644266546
step: 320, loss: 0.04695873335003853
step: 330, loss: 0.0016008830862119794
step: 340, loss: 0.00016153673641383648
step: 350, loss: 3.007728446391411e-05
step: 360, loss: 0.0035385186783969402
step: 370, loss: 0.0010354931000620127
step: 380, loss: 0.0009354744688607752
step: 390, loss: 0.025006061419844627
step: 400, loss: 0.0009582355269230902
step: 410, loss: 0.003965064883232117
step: 420, loss: 0.002479644725099206
step: 430, loss: 0.0039405315183103085
step: 440, loss: 0.013557096011936665
step: 450, loss: 0.058064695447683334
step: 460, loss: 0.00047090830048546195
step: 470, loss: 0.009984835051000118
step: 480, loss: 0.0001449586998205632
step: 490, loss: 0.017756681889295578
step: 500, loss: 0.010109839960932732
step: 510, loss: 0.03395247831940651
step: 520, loss: 0.03809336572885513
step: 530, loss: 0.011937389150261879
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9534776600644865, f1=0.9511970534069981,

```

```
best_f1=0.9511970534069981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008387536508962512
step: 10, loss: 0.01535588875412941
step: 20, loss: 3.846878098556772e-05
step: 30, loss: 0.00012591654376592487
step: 40, loss: 0.037288159132003784
step: 50, loss: 0.004377678968012333
step: 60, loss: 0.0007786014466546476
step: 70, loss: 0.0019185710698366165
step: 80, loss: 0.0012112284312024713
step: 90, loss: 0.0007398639572784305
step: 100, loss: 0.005899009294807911
step: 110, loss: 0.0007090480648912489
step: 120, loss: 0.0049295187927782536
step: 130, loss: 0.00031995863537304103
step: 140, loss: 0.00037996505852788687
step: 150, loss: 0.025009987875819206
step: 160, loss: 0.00025089786504395306
step: 170, loss: 0.05264339968562126
step: 180, loss: 0.0012016125256195664
step: 190, loss: 0.000840446213260293
step: 200, loss: 7.152181933633983e-05
step: 210, loss: 2.0563000362017192e-05
step: 220, loss: 0.0003829893539659679
step: 230, loss: 0.0004335980920586735
step: 240, loss: 0.00012744580453727394
step: 250, loss: 0.00030938920099288225
step: 260, loss: 1.9441873519099317e-05
step: 270, loss: 0.008032631129026413
step: 280, loss: 2.2853762857266702e-05
step: 290, loss: 3.3166528737638146e-05
step: 300, loss: 1.2911737030663062e-05
step: 310, loss: 0.0002926603483501822
step: 320, loss: 1.9218450688640587e-05
step: 330, loss: 0.00014466582797467709
step: 340, loss: 0.004494049120694399
step: 350, loss: 2.2533808078151196e-05
step: 360, loss: 0.13149797916412354
step: 370, loss: 0.013267243281006813
step: 380, loss: 9.463900642003864e-05
step: 390, loss: 0.001533180708065629
step: 400, loss: 0.002020918997004628
step: 410, loss: 0.0002888863964471966
step: 420, loss: 6.98176518199034e-05
step: 430, loss: 6.364515138557181e-05
```

```

step: 440, loss: 0.0011956070084124804
step: 450, loss: 0.004891471937298775
step: 460, loss: 0.0042276037856936455
step: 470, loss: 0.0010323625756427646
step: 480, loss: 1.545980376249645e-05
step: 490, loss: 0.000291741598630324
step: 500, loss: 0.0001371951657347381
step: 510, loss: 0.0002768976555671543
step: 520, loss: 0.0001119363441830501
step: 530, loss: 1.8730466763372533e-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=0.9498850574712644, f1=0.9489655172413793,
best_f1=0.9511970534069981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004788925580214709
step: 10, loss: 0.00045720295747742057
step: 20, loss: 0.00038951425813138485
step: 30, loss: 0.00035778668825514615
step: 40, loss: 0.0002718743053264916
step: 50, loss: 0.0013421676121652126
step: 60, loss: 6.247564306249842e-05
step: 70, loss: 0.0005583098973147571
step: 80, loss: 0.0005688766250386834
step: 90, loss: 3.861384902847931e-05
step: 100, loss: 0.005153970792889595
step: 110, loss: 0.0006475679692812264
step: 120, loss: 3.451504744589329e-05
step: 130, loss: 4.618774983100593e-05
step: 140, loss: 0.0015501058660447598
step: 150, loss: 0.0010978425852954388
step: 160, loss: 0.00035869935527443886
step: 170, loss: 0.0002990478533320129
step: 180, loss: 0.0009716037893667817
step: 190, loss: 0.003978434484452009
step: 200, loss: 1.3179848792788107e-05
step: 210, loss: 0.0011901275720447302
step: 220, loss: 3.193335942341946e-05
step: 230, loss: 1.7728341845213436e-05
step: 240, loss: 0.001166121568530798
step: 250, loss: 0.0018146586371585727
step: 260, loss: 0.03128990903496742
step: 270, loss: 0.0013252721400931478
step: 280, loss: 0.00035933509934693575
step: 290, loss: 0.057222384959459305

```

```

step: 300, loss: 6.686628330498934e-05
step: 310, loss: 0.003379667876288295
step: 320, loss: 3.1297546229325235e-05
step: 330, loss: 0.002458003582432866
step: 340, loss: 0.0015019390266388655
step: 350, loss: 0.00017619914433453232
step: 360, loss: 0.0009676216286607087
step: 370, loss: 0.02257116325199604
step: 380, loss: 0.0033422093838453293
step: 390, loss: 0.0008998664561659098
step: 400, loss: 0.0014131562784314156
step: 410, loss: 0.1154732033610344
step: 420, loss: 0.000313865311909467
step: 430, loss: 0.0011656471760943532
step: 440, loss: 0.0015121996402740479
step: 450, loss: 0.00021909204951953143
step: 460, loss: 0.04756340757012367
step: 470, loss: 0.0009847673354670405
step: 480, loss: 0.0021304164547473192
step: 490, loss: 0.004731038119643927
step: 500, loss: 0.00204260740429163
step: 510, loss: 0.01594516821205616
step: 520, loss: 0.0037991940043866634
step: 530, loss: 0.0029848816338926554
/usr/local/lib/python3.7/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=0.9513314967860421, f1=0.9517241379310345,
best_f1=0.9511970534069981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0048949215561151505
step: 10, loss: 0.001749546849168837
step: 20, loss: 0.002386461477726698
step: 30, loss: 0.00163445551879704
step: 40, loss: 0.00038417376345023513
step: 50, loss: 0.0014362952206283808
step: 60, loss: 0.01765322871506214
step: 70, loss: 0.00026242336025461555
step: 80, loss: 0.00028451718389987946
step: 90, loss: 0.0006479000439867377
step: 100, loss: 2.577073610154912e-05
step: 110, loss: 0.00198144163005054
step: 120, loss: 3.0102766686468385e-05
step: 130, loss: 0.00022094370797276497
step: 140, loss: 8.707470988156274e-05
step: 150, loss: 0.0012557778973132372

```



```

step: 160, loss: 2.016092730627861e-05
step: 170, loss: 0.00014128776092547923
step: 180, loss: 0.0035509001463651657
step: 190, loss: 0.002065237844362855
step: 200, loss: 0.012677205726504326
step: 210, loss: 0.0006308676092885435
step: 220, loss: 3.9999802538659424e-05
step: 230, loss: 0.0011719365138560534
step: 240, loss: 0.0007140521192923188
step: 250, loss: 7.29131861589849e-05
step: 260, loss: 8.44935275381431e-05
step: 270, loss: 8.69157156557776e-05
step: 280, loss: 2.5778192139114253e-05
step: 290, loss: 0.0007039143238216639
step: 300, loss: 5.647767102345824e-05
step: 310, loss: 0.0018607955425977707
step: 320, loss: 9.713815961731598e-05
step: 330, loss: 0.00044830917613580823
step: 340, loss: 0.0013933827867731452
step: 350, loss: 0.001604508957825601
step: 360, loss: 8.187563071260229e-05
step: 370, loss: 0.017392661422491074
step: 380, loss: 2.1505231416085735e-05
step: 390, loss: 5.7013879995793104e-05
step: 400, loss: 0.027454880997538567
step: 410, loss: 0.0004026635142508894
step: 420, loss: 0.0010624720016494393
step: 430, loss: 2.0749341274495237e-05
step: 440, loss: 0.000998315284959972
step: 450, loss: 0.001605934463441372
step: 460, loss: 0.0070405746810138226
step: 470, loss: 3.320078758406453e-05
step: 480, loss: 0.002150064567103982
step: 490, loss: 0.0012870091013610363
step: 500, loss: 0.0049903844483196735
step: 510, loss: 1.65438286785502e-05
step: 520, loss: 2.8943659344804473e-05
step: 530, loss: 0.011735543608665466
/usr/local/lib/python3.7/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=0.9502510269283433, f1=0.9488117001828152,
best_f1=0.9511970534069981
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:23, 242.12it/s]
load_f1 = 0.951411383618695
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 [00:22, 193.98it/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: 0.4554867148399353
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=0.288659793814433, f1=0.2828282828282828,
best_f1=0.2828282828282828
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.488212525844574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2947368421052632, f1=0.288659793814433,
best_f1=0.288659793814433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42301279306411743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3888888888888889, f1=0.3181818181818182,
best_f1=0.3181818181818182

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25951653718948364
/usr/local/lib/python3.7/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=0.35443037974683544, f1=0.31168831168831174,
best_f1=0.3181818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35635972023010254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.46153846153846156, f1=0.423076923076923,
best_f1=0.423076923076923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2818308472633362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4827586206896552, f1=0.40740740740740744,
best_f1=0.40740740740740744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37723979353904724
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=0.5, 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: 0.2973957657814026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, f1=0.588235294117647,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1755973845720291
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.8275862068965518, f1=0.588235294117647,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.234860360622406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.896551724137931, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14167211949825287
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9285714285714286, f1=0.588235294117647,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15590806305408478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9285714285714286, f1=0.6451612903225806,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021982619538903236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9285714285714286, f1=0.6470588235294117,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014075307175517082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9285714285714286, f1=0.6060606060606061,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05630442127585411
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9285714285714286, f1=0.6060606060606061,
best_f1=0.588235294117647
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             : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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, 131659.77it/s]
/usr/local/lib/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 = 0.9333333333333333
real_f1 = 0.9032258064516129
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, 199.81it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

```



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

```
# Step 3: run the matcher
```

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

```
Downloading: 100% 481/481 [00:00<00:00, 420kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 657kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 500kB/s]
Downloading: 100% 501M/501M [00:09<00:00, 54.6MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.
```

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

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

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

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

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

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

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

```
step: 0, loss: 0.5988742709159851
```

```
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: 0.44877195358276367
step: 20, loss: 0.5750805139541626
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.3803631663322449
step: 40, loss: 0.3399450480937958
step: 50, loss: 0.5401262640953064
step: 60, loss: 0.510035514831543
step: 70, loss: 0.4489377737045288
step: 80, loss: 0.6608836054801941
step: 90, loss: 0.44442063570022583
step: 100, loss: 0.4788627624511719
step: 110, loss: 0.6273500323295593
step: 120, loss: 0.4224046468734741
step: 130, loss: 0.32672834396362305
step: 140, loss: 0.3960535526275635
step: 150, loss: 0.5974320769309998
step: 160, loss: 0.4828259348869324
step: 170, loss: 0.4810259938240051
step: 180, loss: 0.4071228802204132
step: 190, loss: 0.27786150574684143
step: 200, loss: 0.4586041271686554
step: 210, loss: 0.35774755477905273
step: 220, loss: 0.6720524430274963
step: 230, loss: 0.3290013074874878
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6719817767653758, f1=0.6769570011025358,
best_f1=0.6769570011025358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14577047526836395
step: 10, loss: 0.2423820197582245
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 20, loss: 0.0848771408200264
step: 30, loss: 0.16732357442378998
step: 40, loss: 0.1644703447818756
step: 50, loss: 0.0668817088007927
step: 60, loss: 0.01605674996972084
step: 70, loss: 0.04050372913479805
step: 80, loss: 0.016732854768633842
step: 90, loss: 0.030768798664212227
step: 100, loss: 0.05824896693229675

```

```

step: 110, loss: 0.011338984593749046
step: 120, loss: 0.12116611748933792
step: 130, loss: 0.04750597104430199
step: 140, loss: 0.0015548896044492722
step: 150, loss: 0.2323436141014099
step: 160, loss: 0.021208543330430984
step: 170, loss: 0.0034544230438768864
step: 180, loss: 0.017104128375649452
step: 190, loss: 0.06510470807552338
step: 200, loss: 0.09330996870994568
step: 210, loss: 0.0640120878815651
step: 220, loss: 0.0017818817868828773
step: 230, loss: 0.005516071803867817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9692982456140351, f1=0.9681668496158068,
best_f1=0.9681668496158068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020574240013957024
step: 10, loss: 0.008094857446849346
step: 20, loss: 0.10125502943992615
step: 30, loss: 0.012914711609482765
step: 40, loss: 0.02685820870101452
step: 50, loss: 0.021541424095630646
step: 60, loss: 0.006838930305093527
step: 70, loss: 0.004257041960954666
step: 80, loss: 0.003911340609192848
step: 90, loss: 0.046334005892276764
step: 100, loss: 0.02781560644507408
step: 110, loss: 0.004326649010181427
step: 120, loss: 0.017373302951455116
step: 130, loss: 0.059622593224048615
step: 140, loss: 0.0012924508191645145
step: 150, loss: 0.07773029059171677
step: 160, loss: 0.011578748002648354
step: 170, loss: 0.0026510150637477636
step: 180, loss: 0.009014284238219261
step: 190, loss: 0.06883082538843155
step: 200, loss: 0.04274752363562584
step: 210, loss: 0.002093641785904765
step: 220, loss: 0.005733908619731665

```

```

step: 230, loss: 0.025927739217877388
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9778270509977827, 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: 0.02081255614757538
step: 10, loss: 0.005025393329560757
step: 20, loss: 0.0038335155695676804
step: 30, loss: 0.001482561812736094
step: 40, loss: 0.11770523339509964
step: 50, loss: 0.00675854180008173
step: 60, loss: 0.01458288636058569
step: 70, loss: 0.02453717403113842
step: 80, loss: 0.004605996422469616
step: 90, loss: 0.007566135376691818
step: 100, loss: 0.004277362953871489
step: 110, loss: 0.0007282656151801348
step: 120, loss: 0.010624973103404045
step: 130, loss: 0.0016297310357913375
step: 140, loss: 0.009675806388258934
step: 150, loss: 0.012149486690759659
step: 160, loss: 0.0006700063240714371
step: 170, loss: 0.0007152763428166509
step: 180, loss: 0.10467330366373062
step: 190, loss: 0.008312913589179516
step: 200, loss: 0.01271775271743536
step: 210, loss: 0.0005046485457569361
step: 220, loss: 0.00034547605901025236
step: 230, loss: 0.00124348362442106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9810479375696767, f1=0.9832026875699889,
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: 0.03343254700303078
step: 10, loss: 0.002158974064514041
step: 20, loss: 0.011700446717441082
step: 30, loss: 0.0006957028526812792
step: 40, loss: 0.002358936471864581
step: 50, loss: 0.01987260952591896
step: 60, loss: 0.01016218587756157
step: 70, loss: 0.0019604219123721123
step: 80, loss: 0.12432009726762772
step: 90, loss: 0.020012738183140755
step: 100, loss: 0.0004735897236969322
step: 110, loss: 0.0021300874650478363
step: 120, loss: 0.00142768956720829
step: 130, loss: 0.00028896654839627445
step: 140, loss: 0.0005584970349445939
step: 150, loss: 0.04345701262354851
step: 160, loss: 0.0011923739220947027
step: 170, loss: 0.04489569365978241
step: 180, loss: 0.004416155628859997
step: 190, loss: 0.003486737608909607
step: 200, loss: 0.01171891950070858
step: 210, loss: 0.0057112895883619785
step: 220, loss: 0.0007972806924954057
step: 230, loss: 0.013806146569550037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9865470852017937, f1=0.9876265466816648,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029429170535877347
step: 10, loss: 0.012224963866174221
step: 20, loss: 0.0008034572238102555
step: 30, loss: 0.0010421209735795856
step: 40, loss: 0.00048098486149683595
step: 50, loss: 0.0006743766716681421
step: 60, loss: 0.0020649221260100603
step: 70, loss: 0.004915488418191671
step: 80, loss: 0.012685244902968407
step: 90, loss: 0.006530147977173328
step: 100, loss: 0.006985160522162914

```

```

step: 110, loss: 0.02668544463813305
step: 120, loss: 0.0005778195336461067
step: 130, loss: 0.0008453646441921592
step: 140, loss: 0.0006056606653146446
step: 150, loss: 0.00011852899478981271
step: 160, loss: 0.009583380073308945
step: 170, loss: 0.0006725536659359932
step: 180, loss: 0.00015121336036827415
step: 190, loss: 0.0011048588203266263
step: 200, loss: 0.0006981715559959412
step: 210, loss: 0.005285822786390781
step: 220, loss: 0.008011643774807453
step: 230, loss: 0.0066483817063272
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.977116704805492,
best_f1=0.9876265466816648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028894897550344467
step: 10, loss: 0.005790805444121361
step: 20, loss: 0.0014086910523474216
step: 30, loss: 0.002363751409575343
step: 40, loss: 0.006668664515018463
step: 50, loss: 0.0013828701339662075
step: 60, loss: 0.000602822401560843
step: 70, loss: 0.00036432623164728284
step: 80, loss: 0.0005645659402944148
step: 90, loss: 0.016721049323678017
step: 100, loss: 0.0015314089832827449
step: 110, loss: 0.0012691137380897999
step: 120, loss: 0.0007764686015434563
step: 130, loss: 0.0010232329368591309
step: 140, loss: 0.0007702879374846816
step: 150, loss: 0.004932239186018705
step: 160, loss: 0.0005135309766046703
step: 170, loss: 0.0007548982976004481
step: 180, loss: 0.0008875152561813593
step: 190, loss: 0.0028712283819913864
step: 200, loss: 0.0014103405410423875
step: 210, loss: 0.010951962321996689
step: 220, loss: 0.0002690760884433985
step: 230, loss: 0.001726479851640761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875424688561721, f1=0.9863945578231292,
best_f1=0.9863945578231292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026647886261343956
step: 10, loss: 0.0053093042224645615
step: 20, loss: 0.0019395974231883883
step: 30, loss: 0.0009240734507329762
step: 40, loss: 0.0010949796997010708
step: 50, loss: 0.003303396049886942
step: 60, loss: 0.0006983595085330307
step: 70, loss: 0.0002531800128053874
step: 80, loss: 0.04656234383583069
step: 90, loss: 0.0003120647161267698
step: 100, loss: 0.0006902713794261217
step: 110, loss: 0.004501631483435631
step: 120, loss: 0.004051173571497202
step: 130, loss: 0.0009668765123933554
step: 140, loss: 0.00041698256973177195
step: 150, loss: 0.045771121978759766
step: 160, loss: 0.004339843522757292
step: 170, loss: 0.0009010470821522176
step: 180, loss: 0.00041116346255876124
step: 190, loss: 0.001498334459029138
step: 200, loss: 0.1339474469423294
step: 210, loss: 0.00941880326718092
step: 220, loss: 0.0024009868502616882
step: 230, loss: 0.0006676848279312253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.987709497206704, f1=0.9832402234636871,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013465315569192171
step: 10, loss: 0.0009987428784370422
step: 20, loss: 0.02660791203379631

```

```

step: 30, loss: 0.006987003609538078
step: 40, loss: 0.0009467501076869667
step: 50, loss: 0.0036808711010962725
step: 60, loss: 0.0003035548434127122
step: 70, loss: 0.046159472316503525
step: 80, loss: 0.0006264628609642386
step: 90, loss: 0.03216751292347908
step: 100, loss: 0.00044467722182162106
step: 110, loss: 0.00025326095055788755
step: 120, loss: 0.032798897475004196
step: 130, loss: 0.00040042263572104275
step: 140, loss: 0.000485033611766994
step: 150, loss: 0.0012439253041520715
step: 160, loss: 0.00088405457790941
step: 170, loss: 0.00024887945619411767
step: 180, loss: 0.00031730643240734935
step: 190, loss: 0.00024192468845285475
step: 200, loss: 0.0007224529981613159
step: 210, loss: 0.0008660555467940867
step: 220, loss: 0.0013245546724647284
step: 230, loss: 0.0003215586766600609
/usr/local/lib/python3.7/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=0.9777777777777777, f1=0.9764837625979844,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005582178710028529
step: 10, loss: 0.0008983959560282528
step: 20, loss: 0.001045369659550488
step: 30, loss: 0.0005955339292995632
step: 40, loss: 0.003477913560345769
step: 50, loss: 0.00044181282282806933
step: 60, loss: 0.001218644087202847
step: 70, loss: 0.006352212280035019
step: 80, loss: 0.001136494567617774
step: 90, loss: 0.002485633594915271
step: 100, loss: 0.00047448568511754274
step: 110, loss: 0.005789379123598337
step: 120, loss: 0.000498503795824945
step: 130, loss: 0.0014894182095304132
step: 140, loss: 0.0020482996478676796
step: 150, loss: 0.004158106632530689
step: 160, loss: 0.00019135934417136014
step: 170, loss: 0.0003376446838956326
step: 180, loss: 0.001471665222197771

```



```

step: 190, loss: 0.000616195669863373
step: 200, loss: 0.0008925217553041875
step: 210, loss: 0.0008471020846627653
step: 220, loss: 0.014808552339673042
step: 230, loss: 0.00032557491795159876
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9865771812080537, f1=0.9864559819413092,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003028957871720195
step: 10, loss: 0.00050493108574301
step: 20, loss: 0.000445204961579293
step: 30, loss: 0.000207044868147932
step: 40, loss: 9.990005491999909e-05
step: 50, loss: 0.00019608737784437835
step: 60, loss: 0.009477490559220314
step: 70, loss: 0.00015817700477782637
step: 80, loss: 0.00019246494048275054
step: 90, loss: 0.10002193599939346
step: 100, loss: 0.00734030082821846
step: 110, loss: 0.0006291450699791312
step: 120, loss: 0.0006115615251474082
step: 130, loss: 0.00031338122789748013
step: 140, loss: 0.005071649793535471
step: 150, loss: 0.0005817816709168255
step: 160, loss: 0.02956586703658104
step: 170, loss: 0.02421567402780056
step: 180, loss: 0.0012982703046873212
step: 190, loss: 0.0011772912694141269
step: 200, loss: 0.007726127747446299
step: 210, loss: 0.00047076825285330415
step: 220, loss: 0.0004677014658227563
step: 230, loss: 0.0005583917954936624
/usr/local/lib/python3.7/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=0.9863945578231292, f1=0.9852440408626559,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005297252791933715
step: 10, loss: 0.0001894853194244206
step: 20, loss: 0.008836890570819378

```

```

step: 30, loss: 0.030184857547283173
step: 40, loss: 0.0006914843106642365
step: 50, loss: 0.00078981724800542
step: 60, loss: 0.004162334371358156
step: 70, loss: 0.0005335782188922167
step: 80, loss: 0.0001612827618373558
step: 90, loss: 0.00502042518928647
step: 100, loss: 0.00022866338258609176
step: 110, loss: 0.00022779408027417958
step: 120, loss: 0.00022560660727322102
step: 130, loss: 0.00016373617108911276
step: 140, loss: 0.0009761280380189419
step: 150, loss: 0.0008691304246895015
step: 160, loss: 0.011182989925146103
step: 170, loss: 0.0010565688135102391
step: 180, loss: 0.00016174939810298383
step: 190, loss: 0.010999809019267559
step: 200, loss: 0.0001856762100942433
step: 210, loss: 0.03310490399599075
step: 220, loss: 0.013811103999614716
step: 230, loss: 0.0058330995962023735
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9876543209876544, f1=0.9876265466816648,
best_f1=0.9832402234636871
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00047344027552753687
step: 10, loss: 0.0009559283498674631
step: 20, loss: 0.0007698938134126365
step: 30, loss: 0.00045534520177170634
step: 40, loss: 0.0011872192844748497
step: 50, loss: 0.0076684495434165
step: 60, loss: 0.00025093386648222804
step: 70, loss: 0.0003506752254907042
step: 80, loss: 0.0003124817449133843
step: 90, loss: 0.0002273710852023214
step: 100, loss: 0.0012693156022578478
step: 110, loss: 0.05463137477636337
step: 120, loss: 0.00033352914033457637
step: 130, loss: 0.0002637551224324852
step: 140, loss: 0.0002540447167120874
step: 150, loss: 0.00016122199303936213
step: 160, loss: 0.0024153655394911766
step: 170, loss: 0.0010995124466717243
step: 180, loss: 0.03213297575712204

```

```

step: 190, loss: 0.0004245094023644924
step: 200, loss: 3.433601523283869e-05
step: 210, loss: 0.00248124310746789
step: 220, loss: 0.0002470972831360996
step: 230, loss: 0.00017160651623271406
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887640449438202, 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: 0.0001186999143101275
step: 10, loss: 0.0001945483236340806
step: 20, loss: 0.000150268358993344
step: 30, loss: 0.00016832136316224933
step: 40, loss: 0.0001337233989033848
step: 50, loss: 6.382591527653858e-05
step: 60, loss: 0.00011307279055472463
step: 70, loss: 0.00016689859330654144
step: 80, loss: 0.00013394805137068033
step: 90, loss: 0.0018677872139960527
step: 100, loss: 0.00014659845328424126
step: 110, loss: 0.00014639945584349334
step: 120, loss: 3.691762321977876e-05
step: 130, loss: 0.00028863013722002506
step: 140, loss: 0.00010298852430423722
step: 150, loss: 4.6965466026449576e-05
step: 160, loss: 0.00018994529091287404
step: 170, loss: 0.00010035693412646651
step: 180, loss: 0.00013823984772898257
step: 190, loss: 9.427704935660586e-05
step: 200, loss: 0.00018147194350603968
step: 210, loss: 0.00016051468264777213
step: 220, loss: 0.00018751125026028603
step: 230, loss: 9.187760588247329e-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=0.9887640449438202, f1=0.987598647125141,
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: 0.019472207874059677
step: 10, loss: 9.64650243986398e-05
step: 20, loss: 0.006692972034215927
step: 30, loss: 0.00018541593453846872
step: 40, loss: 6.508832302642986e-05
step: 50, loss: 0.00012061780580552295
step: 60, loss: 0.037942878901958466
step: 70, loss: 0.001422895584255457
step: 80, loss: 0.00017217782442457974
step: 90, loss: 6.150559056550264e-05
step: 100, loss: 7.284784805960953e-05
step: 110, loss: 0.00015583913773298264
step: 120, loss: 0.016689244657754898
step: 130, loss: 7.223455759231001e-05
step: 140, loss: 0.005086576100438833
step: 150, loss: 0.0020597586408257484
step: 160, loss: 0.02591414377093315
step: 170, loss: 9.251887968275696e-05
step: 180, loss: 0.00020508210582192987
step: 190, loss: 9.791010961635038e-05
step: 200, loss: 0.0036900744307786226
step: 210, loss: 0.027467917650938034
step: 220, loss: 0.00013347642379812896
step: 230, loss: 0.00011399495997466147
/usr/local/lib/python3.7/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=0.9887640449438202, f1=0.987598647125141,
best_f1=0.9864864864864865
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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, 145.63it/s]
load_f1 = 0.9864864864864865
real_f1 = 0.9865168539325843
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, 130.42it/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: 0.6294487118721008
```

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: 0.46084344387054443
```

```
step: 20, loss: 0.3010409474372864
```

```
step: 30, loss: 0.28471073508262634
```

```
step: 40, loss: 0.1680634319782257
```

```
step: 50, loss: 0.1506551206111908
```

```
step: 60, loss: 0.2902781069278717
```

```
step: 70, loss: 0.2123384028673172
```

```
step: 80, loss: 0.2071753293275833
```

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

```
step: 90, loss: 0.2576635181903839
```

```
step: 100, loss: 0.22529476881027222
```

```
step: 110, loss: 0.19561541080474854
```

```
step: 120, loss: 0.18520404398441315
```

```

step: 130, loss: 0.10581626743078232
step: 140, loss: 0.2831640839576721
step: 150, loss: 0.1169738918542862
step: 160, loss: 0.11334899812936783
step: 170, loss: 0.03547261282801628
step: 180, loss: 0.10232581943273544
step: 190, loss: 0.09289273619651794
step: 200, loss: 0.054073162376880646
step: 210, loss: 0.07472054660320282
step: 220, loss: 0.15820716321468353
step: 230, loss: 0.26092156767845154
step: 240, loss: 0.08194289356470108
step: 250, loss: 0.05285811424255371
step: 260, loss: 0.13941097259521484
step: 270, loss: 0.31958457827568054
step: 280, loss: 0.03595774993300438
step: 290, loss: 0.13231021165847778
step: 300, loss: 0.06853923946619034
step: 310, loss: 0.04517388343811035
step: 320, loss: 0.29053887724876404
step: 330, loss: 0.07823435217142105
step: 340, loss: 0.42000311613082886
step: 350, loss: 0.0965183675289154
step: 360, loss: 0.0780305564403534
step: 370, loss: 0.02882467396557331
step: 380, loss: 0.07596201449632645
step: 390, loss: 0.025159625336527824
step: 400, loss: 0.02756839618086815
step: 410, loss: 0.24726666510105133
step: 420, loss: 0.05806853622198105
step: 430, loss: 0.030859321355819702
step: 440, loss: 0.09228477627038956
step: 450, loss: 0.04609176889061928
step: 460, loss: 0.009975266642868519
step: 470, loss: 0.024591544643044472
step: 480, loss: 0.1481398493051529
step: 490, loss: 0.1569262593984604
step: 500, loss: 0.03385060653090477
step: 510, loss: 0.038959700614213943
step: 520, loss: 0.06709912419319153
step: 530, loss: 0.12726271152496338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9435897435897437, f1=0.9391385767790261,  
best_f1=0.9391385767790261  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.014370106160640717  
step: 10, loss: 0.032308466732501984  
step: 20, loss: 0.01973898336291313  
step: 30, loss: 0.26889580488204956  
step: 40, loss: 0.05777162313461304  
step: 50, loss: 0.08767792582511902  
step: 60, loss: 0.0050775655545294285  
step: 70, loss: 0.019397152587771416  
step: 80, loss: 0.050614774227142334  
step: 90, loss: 0.008489751257002354  
step: 100, loss: 0.06438242644071579  
step: 110, loss: 0.014140615239739418  
step: 120, loss: 0.04552634432911873  
step: 130, loss: 0.015749646350741386  
step: 140, loss: 0.08336767554283142  
step: 150, loss: 0.015177397057414055  
step: 160, loss: 0.024531524628400803  
step: 170, loss: 0.013917814940214157  
step: 180, loss: 0.013570157811045647  
step: 190, loss: 0.03364447131752968  
step: 200, loss: 0.19365862011909485  
step: 210, loss: 0.048728540539741516  
step: 220, loss: 0.0012209188425913453  
step: 230, loss: 0.019363686442375183  
step: 240, loss: 0.00391251128166914  
step: 250, loss: 0.013824624940752983  
step: 260, loss: 0.04369835555553436  
step: 270, loss: 0.03430088236927986  
step: 280, loss: 0.10120926797389984  
step: 290, loss: 0.035153962671756744  
step: 300, loss: 0.025083888322114944  
step: 310, loss: 0.10628688335418701  
step: 320, loss: 0.10141529887914658  
step: 330, loss: 0.1586192399263382  
step: 340, loss: 0.02575584501028061  
step: 350, loss: 0.0017131012864410877  
step: 360, loss: 0.11123308539390564  
step: 370, loss: 0.0025203160475939512  
step: 380, loss: 0.206472247838974  
step: 390, loss: 0.015069154091179371  
step: 400, loss: 0.044863417744636536  
step: 410, loss: 0.04589756205677986  
step: 420, loss: 0.0702364593744278
```



```

step: 430, loss: 0.043070778250694275
step: 440, loss: 0.007018618285655975
step: 450, loss: 0.03525596484541893
step: 460, loss: 0.0391480028629303
step: 470, loss: 0.04130278900265694
step: 480, loss: 0.012964766472578049
step: 490, loss: 0.05336989089846611
step: 500, loss: 0.003346954472362995
step: 510, loss: 0.017810294404625893
step: 520, loss: 0.2508140802383423
step: 530, loss: 0.11831041425466537
/usr/local/lib/python3.7/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=0.9386814200092208, f1=0.9312072892938498,
best_f1=0.9391385767790261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15616920590400696
step: 10, loss: 0.045449815690517426
step: 20, loss: 0.01475432701408863
step: 30, loss: 0.047953471541404724
step: 40, loss: 0.048263344913721085
step: 50, loss: 0.0038347167428582907
step: 60, loss: 0.008374172262847424
step: 70, loss: 0.0015256516635417938
step: 80, loss: 0.04301612451672554
step: 90, loss: 0.08027835190296173
step: 100, loss: 0.05877475067973137
step: 110, loss: 0.018208546563982964
step: 120, loss: 0.1713516265153885
step: 130, loss: 0.0011476392392069101
step: 140, loss: 0.002186114201322198
step: 150, loss: 0.015466188080608845
step: 160, loss: 0.09443355351686478
step: 170, loss: 0.049142010509967804
step: 180, loss: 0.01564883068203926
step: 190, loss: 0.0017493086634203792
step: 200, loss: 0.021657763049006462
step: 210, loss: 0.0152366878464818
step: 220, loss: 0.10292709618806839
step: 230, loss: 0.018314342945814133
step: 240, loss: 0.005830278620123863
step: 250, loss: 0.1849433183670044
step: 260, loss: 0.12042681127786636
step: 270, loss: 0.005229776259511709
step: 280, loss: 0.0028741632122546434

```

```

step: 290, loss: 0.011160937137901783
step: 300, loss: 0.17363010346889496
step: 310, loss: 0.13145290315151215
step: 320, loss: 0.026832222938537598
step: 330, loss: 0.03041400946676731
step: 340, loss: 0.012830483727157116
step: 350, loss: 0.07363753020763397
step: 360, loss: 0.008379733189940453
step: 370, loss: 0.03561396896839142
step: 380, loss: 0.0051995511166751385
step: 390, loss: 0.03558827191591263
step: 400, loss: 0.08756611496210098
step: 410, loss: 0.06326552480459213
step: 420, loss: 0.032239772379398346
step: 430, loss: 0.016122102737426758
step: 440, loss: 0.11718375980854034
step: 450, loss: 0.08332262933254242
step: 460, loss: 0.21443767845630646
step: 470, loss: 0.0013463314389809966
step: 480, loss: 0.07711129635572433
step: 490, loss: 0.13968601822853088
step: 500, loss: 0.020378408953547478
step: 510, loss: 0.036087095737457275
step: 520, loss: 0.008317740634083748
step: 530, loss: 0.005049257539212704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9517625231910947, f1=0.9363891487371375,
best_f1=0.9363891487371375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018166687339544296
step: 10, loss: 0.09357231110334396
step: 20, loss: 0.13736285269260406
step: 30, loss: 0.0904356986284256
step: 40, loss: 0.007505667395889759
step: 50, loss: 0.03455530107021332
step: 60, loss: 0.0030734112951904535
step: 70, loss: 0.0013627121224999428
step: 80, loss: 0.0020008895080536604
step: 90, loss: 0.08459774404764175
step: 100, loss: 0.0006802955758757889

```

```
step: 110, loss: 0.19801105558872223
step: 120, loss: 0.0012829515617340803
step: 130, loss: 0.17385192215442657
step: 140, loss: 0.038616906851530075
step: 150, loss: 0.004321216605603695
step: 160, loss: 0.01066544372588396
step: 170, loss: 0.01538555882871151
step: 180, loss: 0.09318463504314423
step: 190, loss: 0.14965805411338806
step: 200, loss: 0.12997198104858398
step: 210, loss: 0.019343281164765358
step: 220, loss: 0.002382437000051141
step: 230, loss: 0.0053880708292126656
step: 240, loss: 0.019343333318829536
step: 250, loss: 0.15522977709770203
step: 260, loss: 0.006156864110380411
step: 270, loss: 0.15996529161930084
step: 280, loss: 0.016619602218270302
step: 290, loss: 0.028960788622498512
step: 300, loss: 0.0016292588552460074
step: 310, loss: 0.005643143318593502
step: 320, loss: 0.035816848278045654
step: 330, loss: 0.0064092399552464485
step: 340, loss: 0.010668660514056683
step: 350, loss: 0.05915937200188637
step: 360, loss: 0.022788725793361664
step: 370, loss: 0.0012383536668494344
step: 380, loss: 0.008236394263803959
step: 390, loss: 0.0048014638014137745
step: 400, loss: 0.02851530909538269
step: 410, loss: 0.0020401179790496826
step: 420, loss: 0.01290846150368452
step: 430, loss: 0.0020375882741063833
step: 440, loss: 0.0015934911789372563
step: 450, loss: 0.02990454062819481
step: 460, loss: 0.09365583956241608
step: 470, loss: 0.005309964530169964
step: 480, loss: 0.00769841531291604
step: 490, loss: 0.0030765782576054335
step: 500, loss: 0.06330137699842453
step: 510, loss: 0.04654943570494652
step: 520, loss: 0.010906478390097618
step: 530, loss: 0.1399008333683014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9554022988505746, f1=0.947077772664519,
best_f1=0.947077772664519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020958012901246548
step: 10, loss: 0.0064866384491324425
step: 20, loss: 0.006380622275173664
step: 30, loss: 0.06745436042547226
step: 40, loss: 0.0019357366254553199
step: 50, loss: 0.04293128103017807
step: 60, loss: 0.04844074696302414
step: 70, loss: 0.0035822454374283552
step: 80, loss: 0.0027705729007720947
step: 90, loss: 0.03396863490343094
step: 100, loss: 0.1278313398361206
step: 110, loss: 0.0020534421782940626
step: 120, loss: 0.1582702249288559
step: 130, loss: 0.018228400498628616
step: 140, loss: 0.001044100383296609
step: 150, loss: 0.015093093737959862
step: 160, loss: 0.004501063842326403
step: 170, loss: 0.1168854683637619
step: 180, loss: 0.00572368036955595
step: 190, loss: 0.004448534455150366
step: 200, loss: 0.017027828842401505
step: 210, loss: 0.002619853476062417
step: 220, loss: 0.0004542737442534417
step: 230, loss: 0.0026274300180375576
step: 240, loss: 0.0014063834678381681
step: 250, loss: 0.23763523995876312
step: 260, loss: 0.003761162282899022
step: 270, loss: 0.006474972236901522
step: 280, loss: 0.009360259398818016
step: 290, loss: 0.007339816074818373
step: 300, loss: 0.16058598458766937
step: 310, loss: 0.008112573996186256
step: 320, loss: 0.0076815239153802395
step: 330, loss: 0.016306236386299133
step: 340, loss: 0.011107045225799084
step: 350, loss: 0.0024422165006399155
step: 360, loss: 0.00033850979525595903
step: 370, loss: 0.0002753899316303432
step: 380, loss: 0.00018903211457654834
step: 390, loss: 0.02533065527677536
step: 400, loss: 0.002157011302188039

```

```

step: 410, loss: 0.0625295639038086
step: 420, loss: 0.2795514762401581
step: 430, loss: 0.0070699527859687805
step: 440, loss: 0.0012955594575032592
step: 450, loss: 0.009853888303041458
step: 460, loss: 0.004041434731334448
step: 470, loss: 0.012232379987835884
step: 480, loss: 0.005232418887317181
step: 490, loss: 0.027565831318497658
step: 500, loss: 0.0013644833816215396
step: 510, loss: 0.0018259789794683456
step: 520, loss: 0.01367983128875494
step: 530, loss: 0.00043426614138297737
/usr/local/lib/python3.7/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=0.95004712535344, f1=0.9315589353612167,
best_f1=0.947077772664519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00031423053587786853
step: 10, loss: 0.00044022584916092455
step: 20, loss: 0.00023249849618878216
step: 30, loss: 0.000461849442217499
step: 40, loss: 0.000327174027916044
step: 50, loss: 0.0033208562526851892
step: 60, loss: 0.004154004156589508
step: 70, loss: 0.00035007792757824063
step: 80, loss: 8.710945985512808e-05
step: 90, loss: 0.006338822655379772
step: 100, loss: 0.1868760883808136
step: 110, loss: 0.013122234493494034
step: 120, loss: 0.009089747443795204
step: 130, loss: 0.015244494192302227
step: 140, loss: 0.0013539334759116173
step: 150, loss: 0.0412193164229393
step: 160, loss: 0.014603405259549618
step: 170, loss: 0.015382247976958752
step: 180, loss: 0.0009396490058861673
step: 190, loss: 0.14604274928569794
step: 200, loss: 0.023390939459204674
step: 210, loss: 0.0005611263331957161
step: 220, loss: 0.005570302717387676
step: 230, loss: 0.00952934194356203
step: 240, loss: 0.002203246345743537
step: 250, loss: 0.008369041606783867
step: 260, loss: 0.0011234533740207553

```

```

step: 270, loss: 0.001529222703538835
step: 280, loss: 0.0035608098842203617
step: 290, loss: 0.007930213585495949
step: 300, loss: 0.009411062113940716
step: 310, loss: 0.030855532735586166
step: 320, loss: 9.13554904400371e-05
step: 330, loss: 0.0002069202600978315
step: 340, loss: 0.03595688194036484
step: 350, loss: 0.0013202287955209613
step: 360, loss: 0.34638556838035583
step: 370, loss: 0.0033969348296523094
step: 380, loss: 0.002939222613349557
step: 390, loss: 0.0036123197060078382
step: 400, loss: 0.015520965680480003
step: 410, loss: 0.13222736120224
step: 420, loss: 0.023346157744526863
step: 430, loss: 0.0016268329927697778
step: 440, loss: 0.000719530216883868
step: 450, loss: 0.3484381139278412
step: 460, loss: 0.0044417972676455975
step: 470, loss: 0.003668965306133032
step: 480, loss: 0.006206988822668791
step: 490, loss: 0.00977796595543623
step: 500, loss: 0.00239822780713439
step: 510, loss: 0.1688472032546997
step: 520, loss: 0.000510868732817471
step: 530, loss: 0.04175829142332077
/usr/local/lib/python3.7/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=0.9499072356215214, f1=0.9415041782729804,
best_f1=0.947077772664519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04110307618975639
step: 10, loss: 0.0006811785278841853
step: 20, loss: 0.005705020856112242
step: 30, loss: 0.0065760486759245396
step: 40, loss: 0.005689720157533884
step: 50, loss: 0.0071481759659945965
step: 60, loss: 0.0018228967674076557
step: 70, loss: 0.0009675798355601728
step: 80, loss: 0.00088554248213768
step: 90, loss: 0.0005321280914358795
step: 100, loss: 0.00014251108223106712
step: 110, loss: 6.877041596453637e-05
step: 120, loss: 0.0002076852833852172

```

```

step: 130, loss: 5.9007939853472635e-05
step: 140, loss: 0.00022293154324870557
step: 150, loss: 0.0042108213528990746
step: 160, loss: 0.0012780613033100963
step: 170, loss: 0.010833247564733028
step: 180, loss: 0.00032478809589520097
step: 190, loss: 0.04358396679162979
step: 200, loss: 6.446857150876895e-05
step: 210, loss: 0.0008437301148660481
step: 220, loss: 7.829657988622785e-05
step: 230, loss: 0.000555639446247369
step: 240, loss: 8.443755359621719e-05
step: 250, loss: 0.023897182196378708
step: 260, loss: 0.010737648233771324
step: 270, loss: 0.021149197593331337
step: 280, loss: 0.02265063300728798
step: 290, loss: 0.001423499546945095
step: 300, loss: 0.0002122628065990284
step: 310, loss: 0.00019806345517281443
step: 320, loss: 0.048076651990413666
step: 330, loss: 0.002748564351350069
step: 340, loss: 0.000273004436166957
step: 350, loss: 0.003478492610156536
step: 360, loss: 0.02213343046605587
step: 370, loss: 0.0178631991147995
step: 380, loss: 0.00040073125273920596
step: 390, loss: 0.017933489754796028
step: 400, loss: 0.012440475635230541
step: 410, loss: 0.004602615721523762
step: 420, loss: 0.006894825957715511
step: 430, loss: 0.0010039940243586898
step: 440, loss: 7.398512389045209e-05
step: 450, loss: 0.0013498773332685232
step: 460, loss: 0.012376699596643448
step: 470, loss: 0.012598242610692978
step: 480, loss: 0.002203197218477726
step: 490, loss: 0.0024486021138727665
step: 500, loss: 0.000309645984089002
step: 510, loss: 5.3480984206544235e-05
step: 520, loss: 0.04800423979759216
step: 530, loss: 0.0004712879890576005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9576547231270358, f1=0.9448818897637795,  
best_f1=0.9448818897637795  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0004963668761774898  
step: 10, loss: 0.00038396529271267354  
step: 20, loss: 0.0004912480362690985  
step: 30, loss: 0.0039596776477992535  
step: 40, loss: 0.00020485302957240492  
step: 50, loss: 0.0001705717295408249  
step: 60, loss: 0.00044471805449575186  
step: 70, loss: 0.00037882235483266413  
step: 80, loss: 0.000627700996119529  
step: 90, loss: 0.034849394112825394  
step: 100, loss: 0.00040410584188066423  
step: 110, loss: 0.00024356204085052013  
step: 120, loss: 0.001355944317765534  
step: 130, loss: 5.313752990332432e-05  
step: 140, loss: 2.7838028472615406e-05  
step: 150, loss: 0.00032178981928154826  
step: 160, loss: 3.89384149457328e-05  
step: 170, loss: 0.08296039700508118  
step: 180, loss: 0.0005818394711241126  
step: 190, loss: 0.002777561778202653  
step: 200, loss: 0.017203129827976227  
step: 210, loss: 0.05763046815991402  
step: 220, loss: 3.2390271371696144e-05  
step: 230, loss: 0.024661961942911148  
step: 240, loss: 0.0012999252649024129  
step: 250, loss: 0.028568532317876816  
step: 260, loss: 0.0005415261839516461  
step: 270, loss: 0.010215149261057377  
step: 280, loss: 0.0029135812073946  
step: 290, loss: 0.0001817017764551565  
step: 300, loss: 7.784114859532565e-05  
step: 310, loss: 0.0018488753121346235  
step: 320, loss: 0.0007390844984911382  
step: 330, loss: 0.0001655327796470374  
step: 340, loss: 0.01804829202592373  
step: 350, loss: 0.0007984416442923248  
step: 360, loss: 0.04025653004646301  
step: 370, loss: 0.026598650962114334  
step: 380, loss: 7.227720925584435e-05  
step: 390, loss: 0.005783332046121359  
step: 400, loss: 9.347422019345686e-05  
step: 410, loss: 0.001339319976978004  
step: 420, loss: 0.00047384246136061847
```



```

step: 430, loss: 0.002450118539854884
step: 440, loss: 0.004285302013158798
step: 450, loss: 0.005242371000349522
step: 460, loss: 0.002434064634144306
step: 470, loss: 0.05220726877450943
step: 480, loss: 0.006147262640297413
step: 490, loss: 0.00025706732412800193
step: 500, loss: 0.00030623795464634895
step: 510, loss: 0.001415229169651866
step: 520, loss: 0.0006728244479745626
step: 530, loss: 7.47402518754825e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9595160539785947, f1=0.9476124246638852,
best_f1=0.9476124246638852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004725189879536629
step: 10, loss: 0.00884393509477377
step: 20, loss: 3.758583989110775e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.06440910696983337
step: 40, loss: 0.00013497906911652535
step: 50, loss: 2.7469583073980175e-05
step: 60, loss: 0.00022262723359744996
step: 70, loss: 0.0008455213392153382
step: 80, loss: 0.025385336950421333
step: 90, loss: 0.02820964902639389
step: 100, loss: 0.005404123105108738
step: 110, loss: 0.0029677560087293386
step: 120, loss: 0.0025146801490336657
step: 130, loss: 0.0007258543628267944
step: 140, loss: 0.005265086889266968
step: 150, loss: 0.0011712315026670694
step: 160, loss: 0.01243320107460022
step: 170, loss: 0.005830394569784403
step: 180, loss: 0.0025205325800925493
step: 190, loss: 4.053056181874126e-05
step: 200, loss: 0.0009609098779037595
step: 210, loss: 0.021562961861491203
step: 220, loss: 0.00021169664978515357
step: 230, loss: 7.74581785663031e-05

```

```

step: 240, loss: 0.0010581129463389516
step: 250, loss: 0.00045143626630306244
step: 260, loss: 0.0008588404743932188
step: 270, loss: 0.0002804580144584179
step: 280, loss: 0.06678427010774612
step: 290, loss: 0.00010824472701642662
step: 300, loss: 9.922036406351253e-05
step: 310, loss: 0.12187027931213379
step: 320, loss: 2.2988015189184807e-05
step: 330, loss: 0.0011703558266162872
step: 340, loss: 0.04875892773270607
step: 350, loss: 0.028572803363204002
step: 360, loss: 0.00383273814804852
step: 370, loss: 3.329806713736616e-05
step: 380, loss: 0.0001969387085409835
step: 390, loss: 1.698712003417313e-05
step: 400, loss: 0.1296176016330719
step: 410, loss: 0.0009219197672791779
step: 420, loss: 0.005098750814795494
step: 430, loss: 0.010989413596689701
step: 440, loss: 0.0007720901630818844
step: 450, loss: 0.0010499658528715372
step: 460, loss: 2.731293898250442e-05
step: 470, loss: 1.8473432646715082e-05
step: 480, loss: 2.0596755348378792e-05
step: 490, loss: 0.0015283877728506923
step: 500, loss: 2.360665712330956e-05
step: 510, loss: 3.680280497064814e-05
step: 520, loss: 0.0029607280157506466
step: 530, loss: 0.0002282523491885513
/usr/local/lib/python3.7/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=0.9562383612662941, f1=0.9484440315838365,
best_f1=0.9476124246638852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.204721724614501e-05
step: 10, loss: 0.0013057391624897718
step: 20, loss: 2.5867402655421756e-05
step: 30, loss: 0.0003831881913356483
step: 40, loss: 8.523814176442102e-05
step: 50, loss: 1.937480192282237e-05
step: 60, loss: 0.006862139329314232
step: 70, loss: 2.929379479610361e-05
step: 80, loss: 0.00017284754721913487
step: 90, loss: 2.223549563495908e-05

```

```
step: 100, loss: 0.0007168437005020678
step: 110, loss: 0.006998084485530853
step: 120, loss: 8.519539551343769e-05
step: 130, loss: 2.442159893689677e-05
step: 140, loss: 0.0001587360311532393
step: 150, loss: 0.0031765764579176903
step: 160, loss: 0.032074302434921265
step: 170, loss: 5.806858098367229e-05
step: 180, loss: 0.002363645238801837
step: 190, loss: 1.4215534065442625e-05
step: 200, loss: 5.719233013223857e-05
step: 210, loss: 0.0018053707899525762
step: 220, loss: 0.00034391635563224554
step: 230, loss: 1.9102797523373738e-05
step: 240, loss: 1.909904312924482e-05
step: 250, loss: 0.0005850416491739452
step: 260, loss: 0.000256790459388867
step: 270, loss: 1.8596294466988184e-05
step: 280, loss: 0.07416877150535583
step: 290, loss: 1.615629116713535e-05
step: 300, loss: 0.0045054261572659016
step: 310, loss: 0.02805778756737709
step: 320, loss: 3.230669244658202e-05
step: 330, loss: 9.761566616361961e-05
step: 340, loss: 1.3224648682808038e-05
step: 350, loss: 0.0019346538465470076
step: 360, loss: 0.00011395443289075047
step: 370, loss: 6.137911987025291e-05
step: 380, loss: 0.0002738970797508955
step: 390, loss: 2.0727129594888538e-05
step: 400, loss: 0.0050039817579090595
step: 410, loss: 0.0019307382171973586
step: 420, loss: 4.7842589992797e-05
step: 430, loss: 2.4816063159960322e-05
step: 440, loss: 2.952372597064823e-05
step: 450, loss: 0.05498165264725685
step: 460, loss: 0.00035448733251541853
step: 470, loss: 0.0034504050854593515
step: 480, loss: 0.0002168486244045198
step: 490, loss: 0.013921823352575302
step: 500, loss: 0.002046002773568034
step: 510, loss: 1.880121635622345e-05
step: 520, loss: 0.0019465569639578462
step: 530, loss: 0.0016452256822958589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9618604651162791, f1=0.9493729679516953,
best_f1=0.9493729679516953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

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

```
step: 0, loss: 4.847968011745252e-05
step: 10, loss: 7.502982771256939e-05
step: 20, loss: 0.0011806655675172806
step: 30, loss: 0.0008482636767439544
step: 40, loss: 1.4856274901831057e-05
step: 50, loss: 1.9743527445825748e-05
step: 60, loss: 0.0001425251248292625
step: 70, loss: 6.437386764446273e-05
step: 80, loss: 0.0004363569605629891
step: 90, loss: 0.0007571993046440184
step: 100, loss: 3.8275735278148204e-05
step: 110, loss: 1.1745735719159711e-05
step: 120, loss: 4.7679906856501475e-05
step: 130, loss: 2.6997062377631664e-05
step: 140, loss: 1.2792509551218245e-05
step: 150, loss: 0.26116785407066345
step: 160, loss: 0.00025774844107218087
step: 170, loss: 0.00019065629749093205
step: 180, loss: 0.00016205973224714398
step: 190, loss: 0.0010213343193754554
step: 200, loss: 0.0007281369180418551
step: 210, loss: 0.0027254752349108458
step: 220, loss: 0.0025635792408138514
step: 230, loss: 0.04846208170056343
step: 240, loss: 0.004424428101629019
step: 250, loss: 0.0004530347359832376
step: 260, loss: 0.004179708193987608
step: 270, loss: 0.001332540181465447
step: 280, loss: 0.0016525093233212829
step: 290, loss: 0.00021816023217979819
step: 300, loss: 2.2839154553366825e-05
step: 310, loss: 0.0003788984031416476
step: 320, loss: 9.465506445849314e-05
step: 330, loss: 2.3415705072693527e-05
step: 340, loss: 0.00017856679914984852
step: 350, loss: 0.00017514979117549956
step: 360, loss: 0.00027063675224781036
step: 370, loss: 0.0002783800882752985
step: 380, loss: 0.0008053817437030375
step: 390, loss: 0.0009176813182421029
```

```

step: 400, loss: 0.00019203292322345078
step: 410, loss: 0.0015919639263302088
step: 420, loss: 0.00047042890219017863
step: 430, loss: 0.00324227474629879
step: 440, loss: 0.0006968663656152785
step: 450, loss: 0.0008495866786688566
step: 460, loss: 0.00011534420627867803
step: 470, loss: 0.00043254828779026866
step: 480, loss: 0.0014621126465499401
step: 490, loss: 0.015815449878573418
step: 500, loss: 0.0014864346012473106
step: 510, loss: 0.013922791928052902
step: 520, loss: 0.0008844708791002631
step: 530, loss: 0.02843851037323475
/usr/local/lib/python3.7/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=0.9591836734693878, f1=0.9484440315838365,
best_f1=0.9493729679516953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001907297526486218
step: 10, loss: 0.003744573798030615
step: 20, loss: 0.003151805605739355
step: 30, loss: 0.0006770172622054815
step: 40, loss: 0.0018053120002150536
step: 50, loss: 0.008607003837823868
step: 60, loss: 2.7584650524659082e-05
step: 70, loss: 0.0009828177280724049
step: 80, loss: 1.8298262148164213e-05
step: 90, loss: 0.0005431381869129837
step: 100, loss: 0.006256533786654472
step: 110, loss: 9.146887168753892e-05
step: 120, loss: 0.0007451109122484922
step: 130, loss: 6.854709499748424e-05
step: 140, loss: 1.7389145796187222e-05
step: 150, loss: 9.681940355221741e-06
step: 160, loss: 0.0004075367469340563
step: 170, loss: 1.332141982857138e-05
step: 180, loss: 3.1883591873338446e-05
step: 190, loss: 0.00019924102525692433
step: 200, loss: 0.0027123503386974335
step: 210, loss: 0.004430255852639675
step: 220, loss: 6.254106847336516e-05
step: 230, loss: 9.590449190000072e-05
step: 240, loss: 0.0003258737560827285
step: 250, loss: 0.00014693297271151096

```

```

step: 260, loss: 6.081675019231625e-05
step: 270, loss: 0.002084101550281048
step: 280, loss: 0.0012395854573696852
step: 290, loss: 0.00045513574150390923
step: 300, loss: 0.00012865765893366188
step: 310, loss: 1.645418888074346e-05
step: 320, loss: 0.000463193457107991
step: 330, loss: 0.0006567251984961331
step: 340, loss: 1.9709406842594035e-05
step: 350, loss: 6.97917930665426e-05
step: 360, loss: 8.23564623715356e-05
step: 370, loss: 0.0011256583966314793
step: 380, loss: 5.664609489031136e-05
step: 390, loss: 0.0004568634903989732
step: 400, loss: 2.2464357243734412e-05
step: 410, loss: 0.015296847559511662
step: 420, loss: 2.9671711672563106e-05
step: 430, loss: 0.00027438433608040214
step: 440, loss: 0.0007752199890092015
step: 450, loss: 0.003001909703016281
step: 460, loss: 6.066348942113109e-05
step: 470, loss: 0.0008203968754969537
step: 480, loss: 0.0007035603048279881
step: 490, loss: 0.03744194656610489
step: 500, loss: 0.00025265346630476415
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 510, loss: 0.005058138631284237
step: 520, loss: 0.005141590256243944
step: 530, loss: 0.0002837575739249587
/usr/local/lib/python3.7/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=0.9616982002768805, f1=0.9501385041551247,
best_f1=0.9493729679516953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.99373010522686e-05
step: 10, loss: 0.0008307560347020626
step: 20, loss: 0.0034142108634114265
step: 30, loss: 1.6736794350435957e-05
step: 40, loss: 0.0005498139653354883
step: 50, loss: 0.006296930368989706
step: 60, loss: 0.0002008052688324824
step: 70, loss: 0.0006212005391716957
step: 80, loss: 0.001094632432796061
step: 90, loss: 0.000425730919232592
step: 100, loss: 3.62092541763559e-05

```

```
step: 110, loss: 8.694753887539264e-06
step: 120, loss: 0.018369102850556374
step: 130, loss: 0.00011287884990451857
step: 140, loss: 0.00019460135081317276
step: 150, loss: 0.0006767347804270685
step: 160, loss: 0.00014234700938686728
step: 170, loss: 0.00044387171510607004
step: 180, loss: 1.474055079597747e-05
step: 190, loss: 0.0003225103428121656
step: 200, loss: 4.438596806721762e-05
step: 210, loss: 0.00017024856060743332
step: 220, loss: 0.0013508122647181153
step: 230, loss: 0.007837502285838127
step: 240, loss: 0.00024678066256456077
step: 250, loss: 4.013921716250479e-05
step: 260, loss: 0.0001788333320291713
step: 270, loss: 0.00013421195035334677
step: 280, loss: 7.54317661630921e-05
step: 290, loss: 4.4924749090569094e-05
step: 300, loss: 0.14877746999263763
step: 310, loss: 1.1399226423236541e-05
step: 320, loss: 3.785472654271871e-05
step: 330, loss: 0.0013367636129260063
step: 340, loss: 0.0002201833121944219
step: 350, loss: 7.055662081256742e-06
step: 360, loss: 7.714897219557315e-05
step: 370, loss: 5.5467604397563264e-05
step: 380, loss: 0.0001876993919722736
step: 390, loss: 4.998964141122997e-05
step: 400, loss: 8.011707541299984e-05
step: 410, loss: 4.052004805998877e-05
step: 420, loss: 4.534049367066473e-05
step: 430, loss: 0.002139525953680277
step: 440, loss: 1.4289564205682836e-05
step: 450, loss: 0.0001545609993627295
step: 460, loss: 1.6443151253042743e-05
step: 470, loss: 0.00013630803732667118
step: 480, loss: 3.2956741051748395e-05
step: 490, loss: 1.8245717001263984e-05
step: 500, loss: 8.247715413745027e-06
step: 510, loss: 2.4271048459922895e-05
step: 520, loss: 6.729913002345711e-05
step: 530, loss: 3.690601806738414e-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=0.9595536959553697, f1=0.9460465116279071,
best_f1=0.9493729679516953
```

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

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

step: 0, loss: 6.185098027344793e-05
step: 10, loss: 0.0001687353360466659
step: 20, loss: 7.819337042747065e-06
step: 30, loss: 0.0001351336541119963
step: 40, loss: 8.098726539174095e-06
step: 50, loss: 7.400655886158347e-05
step: 60, loss: 2.228381890745368e-05
step: 70, loss: 3.846337494906038e-05
step: 80, loss: 1.055737811839208e-05
step: 90, loss: 8.281239388452377e-06
step: 100, loss: 1.3410674000624567e-05
step: 110, loss: 0.0005917245871387422
step: 120, loss: 1.0479144293640275e-05
step: 130, loss: 9.003960258269217e-06
step: 140, loss: 8.618509309599176e-05
step: 150, loss: 0.0003493239055387676
step: 160, loss: 0.00012746294669341296
step: 170, loss: 2.2352254745783284e-05
step: 180, loss: 7.15250871508033e-06
step: 190, loss: 1.3213127203925978e-05
step: 200, loss: 1.2691829397226684e-05
step: 210, loss: 8.5867368397885e-06
step: 220, loss: 9.268412213714328e-06
step: 230, loss: 1.6859283277881332e-05
step: 240, loss: 8.951777090260293e-06
step: 250, loss: 0.0003518106241244823
step: 260, loss: 5.567346670432016e-05
step: 270, loss: 3.806131280725822e-05
step: 280, loss: 8.013027581910137e-06
step: 290, loss: 6.325502454274101e-06
step: 300, loss: 1.7094571376219392e-05
step: 310, loss: 8.240265742642805e-06
step: 320, loss: 1.0289155397913419e-05
step: 330, loss: 1.3440346265269909e-05
step: 340, loss: 0.00011226743663428351
step: 350, loss: 8.605355105828494e-06
step: 360, loss: 1.571957182022743e-05
step: 370, loss: 0.0010725281899794936
step: 380, loss: 0.0030251445714384317
step: 390, loss: 0.016505559906363487
step: 400, loss: 0.001736168167553842
step: 410, loss: 6.485697213065578e-06
step: 420, loss: 8.817682100925595e-06
step: 430, loss: 3.3336400520056486e-05
step: 440, loss: 9.041171324497554e-06


```

step: 450, loss: 3.325192301417701e-05
step: 460, loss: 0.026606617495417595
step: 470, loss: 2.387539279880002e-05
step: 480, loss: 5.651135870721191e-05
step: 490, loss: 1.1417838322813623e-05
step: 500, loss: 0.02744501270353794
step: 510, loss: 0.009799026884138584
step: 520, loss: 9.091039100894704e-05
step: 530, loss: 4.845330840907991e-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=0.9598506069094305, f1=0.9474662947466295,
best_f1=0.9493729679516953
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.813636016682722e-05
step: 10, loss: 0.00010613464110065252
step: 20, loss: 0.00029036105843260884
step: 30, loss: 0.005734777078032494
step: 40, loss: 9.909144864650443e-06
step: 50, loss: 1.868775325419847e-05
step: 60, loss: 0.0005592455854639411
step: 70, loss: 0.0005352570442482829
step: 80, loss: 0.0001359269954264164
step: 90, loss: 1.7522135749459267e-05
step: 100, loss: 4.754302790388465e-05
step: 110, loss: 9.02478423085995e-05
step: 120, loss: 1.4076874322199728e-05
step: 130, loss: 9.64837454375811e-06
step: 140, loss: 0.0005152292433194816
step: 150, loss: 7.9422670751228e-06
step: 160, loss: 5.675913052982651e-05
step: 170, loss: 6.4670412029954605e-06
step: 180, loss: 0.0004658540419768542
step: 190, loss: 0.0003017097187694162
step: 200, loss: 0.0014874188927933574
step: 210, loss: 8.009310477063991e-06
step: 220, loss: 8.91452236828627e-06
step: 230, loss: 0.08695335686206818
step: 240, loss: 8.16949068394024e-06
step: 250, loss: 6.638422291871393e-06
step: 260, loss: 7.476613063772675e-06
step: 270, loss: 1.192820582218701e-05
step: 280, loss: 7.409553745674202e-06
step: 290, loss: 0.0017088173190131783
step: 300, loss: 8.292402526421938e-06

```

```

step: 310, loss: 0.045618489384651184
step: 320, loss: 0.0001453185104764998
step: 330, loss: 9.335496542917099e-06
step: 340, loss: 1.2796122973668389e-05
step: 350, loss: 0.0017107271123677492
step: 360, loss: 1.4218582691682968e-05
step: 370, loss: 0.0001351788523606956
step: 380, loss: 0.0001233293442055583
step: 390, loss: 8.355235331691802e-05
step: 400, loss: 0.0036732445005327463
step: 410, loss: 0.029907532036304474
step: 420, loss: 7.927353181003127e-06
step: 430, loss: 6.932710675755516e-06
step: 440, loss: 0.12416501343250275
step: 450, loss: 0.0009838222758844495
step: 460, loss: 1.1201612323930021e-05
step: 470, loss: 7.376034318440361e-06
step: 480, loss: 0.000571750570088625
step: 490, loss: 0.021866418421268463
step: 500, loss: 0.00013217376545071602
step: 510, loss: 2.8266953449929133e-05
step: 520, loss: 9.596209565643221e-06
step: 530, loss: 8.92940897756489e-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=0.9598506069094305, f1=0.9491367242183854,
best_f1=0.9493729679516953
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:21, 262.48it/s]
load_f1 = 0.9613414066138799
real_f1 = 0.9617537313432836
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, 198.65it/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, 311kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 821kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 528kB/s]
Downloading: 100% 501M/501M [00:13<00:00, 36.3MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

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

```
step: 0, loss: 0.5238576531410217
```

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

```
step: 10, loss: 0.4674552083015442
```

```
step: 20, loss: 0.4959307909011841
```

```
step: 30, loss: 0.32108527421951294
```

```

step: 40, loss: 0.30842462182044983
step: 50, loss: 0.4235861301422119
step: 60, loss: 0.451276034116745
step: 70, loss: 0.3041010797023773
step: 80, loss: 0.4141964912414551
step: 90, loss: 0.24703529477119446
step: 100, loss: 0.2530973553657532
step: 110, loss: 0.27413076162338257
step: 120, loss: 0.3717157542705536
step: 130, loss: 0.21942272782325745
step: 140, loss: 0.42579391598701477
step: 150, loss: 0.21918290853500366
step: 160, loss: 0.5042417049407959
step: 170, loss: 0.1830274611711502
step: 180, loss: 0.3036159873008728
step: 190, loss: 0.2936713695526123
step: 200, loss: 0.16739287972450256
step: 210, loss: 0.3181871175765991
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5719557195571957, f1=0.5981308411214954,
best_f1=0.5981308411214954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.125617116689682
step: 10, loss: 0.08944663405418396
step: 20, loss: 0.2687239646911621
step: 30, loss: 0.2119007110595703
step: 40, loss: 0.21672365069389343
step: 50, loss: 0.1257743239402771
step: 60, loss: 0.25200045108795166
step: 70, loss: 0.1909729391336441
step: 80, loss: 0.1698886752128601
step: 90, loss: 0.23072247207164764
step: 100, loss: 0.27546757459640503
step: 110, loss: 0.30231136083602905
step: 120, loss: 0.07476072758436203
step: 130, loss: 0.08749701082706451
step: 140, loss: 0.08331344276666641
step: 150, loss: 0.24629169702529907
step: 160, loss: 0.03278458118438721
step: 170, loss: 0.35875800251960754

```

```

step: 180, loss: 0.14442606270313263
step: 190, loss: 0.28357064723968506
step: 200, loss: 0.10968055576086044
step: 210, loss: 0.0800894945859909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6630236794171219, f1=0.6703703703703704,
best_f1=0.6703703703703704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05433919280767441
step: 10, loss: 0.047176919877529144
step: 20, loss: 0.26643210649490356
step: 30, loss: 0.0691947266459465
step: 40, loss: 0.248758465051651
step: 50, loss: 0.10302668809890747
step: 60, loss: 0.2360609918832779
step: 70, loss: 0.059625543653964996
step: 80, loss: 0.16526947915554047
step: 90, loss: 0.03753424808382988
step: 100, loss: 0.15540120005607605
step: 110, loss: 0.0978262647986412
step: 120, loss: 0.14027157425880432
step: 130, loss: 0.29687830805778503
step: 140, loss: 0.0808335393667221
step: 150, loss: 0.13162410259246826
step: 160, loss: 0.11502773314714432
step: 170, loss: 0.22401022911071777
step: 180, loss: 0.1196407824754715
step: 190, loss: 0.01600201614201069
step: 200, loss: 0.09849867224693298
step: 210, loss: 0.19323548674583435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6776556776556777, f1=0.6816479400749063,
best_f1=0.6816479400749063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03319893777370453
step: 10, loss: 0.10077553987503052
step: 20, loss: 0.060439787805080414
step: 30, loss: 0.07061123102903366
step: 40, loss: 0.03486516699194908
step: 50, loss: 0.09075067192316055
step: 60, loss: 0.13524161279201508
step: 70, loss: 0.03921116515994072
step: 80, loss: 0.07359403371810913
step: 90, loss: 0.04969121143221855
step: 100, loss: 0.18194226920604706
step: 110, loss: 0.4403432011604309
step: 120, loss: 0.139717698097229
step: 130, loss: 0.18161331117153168
step: 140, loss: 0.2496328055858612
step: 150, loss: 0.022324813529849052
step: 160, loss: 0.06244628131389618
step: 170, loss: 0.04545455798506737
step: 180, loss: 0.030737219378352165
step: 190, loss: 0.13565963506698608
step: 200, loss: 0.061110787093639374
step: 210, loss: 0.353935182094574
/usr/local/lib/python3.7/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=0.673773987206823, f1=0.6893617021276596,
best_f1=0.6816479400749063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09151702374219894
step: 10, loss: 0.05910755693912506
step: 20, loss: 0.06430480629205704
step: 30, loss: 0.009805974550545216
step: 40, loss: 0.026088669896125793
step: 50, loss: 0.03681938722729683
step: 60, loss: 0.03226036578416824
step: 70, loss: 0.1759769320487976
step: 80, loss: 0.0905127003788948
step: 90, loss: 0.06204579770565033
step: 100, loss: 0.1623319536447525
step: 110, loss: 0.1071503758430481
step: 120, loss: 0.0641310065984726
step: 130, loss: 0.1005108505487442
step: 140, loss: 0.07183638215065002
step: 150, loss: 0.03479728475213051

```

```

step: 160, loss: 0.01656755618751049
step: 170, loss: 0.10002229362726212
step: 180, loss: 0.07602417469024658
step: 190, loss: 0.10995590686798096
step: 200, loss: 0.061656493693590164
step: 210, loss: 0.12473759800195694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7074569789674952, f1=0.6976744186046512,
best_f1=0.6976744186046512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01611025258898735
step: 10, loss: 0.07416550070047379
step: 20, loss: 0.03204837813973427
step: 30, loss: 0.010543429292738438
step: 40, loss: 0.06577436625957489
step: 50, loss: 0.025563251227140427
step: 60, loss: 0.04792409762740135
step: 70, loss: 0.06946922093629837
step: 80, loss: 0.10355684161186218
step: 90, loss: 0.03777335584163666
step: 100, loss: 0.07929396629333496
step: 110, loss: 0.11027425527572632
step: 120, loss: 0.001169333583675325
step: 130, loss: 0.002236920641735196
step: 140, loss: 0.09044108539819717
step: 150, loss: 0.06294510513544083
step: 160, loss: 0.017835991457104683
step: 170, loss: 0.05548140034079552
step: 180, loss: 0.0395260788500309
step: 190, loss: 0.0280917938798666
step: 200, loss: 0.10044499486684799
step: 210, loss: 0.02791074849665165
/usr/local/lib/python3.7/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=0.6944971537001896, f1=0.7132075471698114,
best_f1=0.6976744186046512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.050717394798994064
step: 10, loss: 0.014379294589161873
step: 20, loss: 0.009558100253343582
step: 30, loss: 0.10501149296760559
step: 40, loss: 0.010893323458731174
step: 50, loss: 0.04162605106830597
step: 60, loss: 0.08663620799779892
step: 70, loss: 0.09363315999507904
step: 80, loss: 0.08499173074960709
step: 90, loss: 0.25853782892227173
step: 100, loss: 0.017028076574206352
step: 110, loss: 0.026043487712740898
step: 120, loss: 0.05020619183778763
step: 130, loss: 0.13189639151096344
step: 140, loss: 0.03493725135922432
step: 150, loss: 0.024144358932971954
step: 160, loss: 0.19621706008911133
step: 170, loss: 0.032059602439403534
step: 180, loss: 0.03199980780482292
step: 190, loss: 0.04380644112825394
step: 200, loss: 0.06881153583526611
step: 210, loss: 0.1962439864873886
/usr/local/lib/python3.7/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=0.7061068702290075, f1=0.688212927756654,
best_f1=0.6976744186046512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07073904573917389
step: 10, loss: 0.01158777717500925
step: 20, loss: 0.020877696573734283
step: 30, loss: 0.058618150651454926
step: 40, loss: 0.01607898809015751
step: 50, loss: 0.0007690550992265344
step: 60, loss: 0.013153715059161186
step: 70, loss: 0.051333073526620865
step: 80, loss: 0.030208751559257507
step: 90, loss: 0.10446573793888092
step: 100, loss: 0.05778978392481804
step: 110, loss: 0.019082654267549515
step: 120, loss: 0.05176859721541405
step: 130, loss: 0.0016944085946306586
step: 140, loss: 0.06779953092336655
step: 150, loss: 0.10450602322816849
step: 160, loss: 0.143738254904747
step: 170, loss: 0.2249627262353897

```

```

step: 180, loss: 0.03359673172235489
step: 190, loss: 0.0658109113574028
step: 200, loss: 0.05559756979346275
step: 210, loss: 0.06794232875108719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7151515151515151, f1=0.7011952191235059,
best_f1=0.7011952191235059
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06446558982133865
step: 10, loss: 0.043955374509096146
step: 20, loss: 0.0315210223197937
step: 30, loss: 0.0006602059584110975
step: 40, loss: 0.026879331097006798
step: 50, loss: 0.01672861911356449
step: 60, loss: 0.015516112558543682
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.012665837071835995
step: 80, loss: 0.0008634625119157135
step: 90, loss: 0.051791854202747345
step: 100, loss: 0.07114269584417343
step: 110, loss: 0.04551207274198532
step: 120, loss: 0.04135250672698021
step: 130, loss: 0.08837839215993881
step: 140, loss: 0.04128584638237953
step: 150, loss: 0.12617523968219757
step: 160, loss: 0.20065324008464813
step: 170, loss: 0.05196583271026611
step: 180, loss: 0.033889032900333405
step: 190, loss: 0.017074506729841232
step: 200, loss: 0.019632842391729355
step: 210, loss: 0.04113771766424179
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7198364008179958, f1=0.7122736418511065,
best_f1=0.7122736418511065

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005031258799135685
step: 10, loss: 0.015708476305007935
step: 20, loss: 0.0009010635549202561
step: 30, loss: 0.0010610633762553334
step: 40, loss: 0.08629107475280762
step: 50, loss: 0.010592254810035229
step: 60, loss: 0.0010572544997557998
step: 70, loss: 0.003651789855211973
step: 80, loss: 0.026093903928995132
step: 90, loss: 0.0727626383304596
step: 100, loss: 0.02765985205769539
step: 110, loss: 0.0008149893255904317
step: 120, loss: 0.04592831805348396
step: 130, loss: 0.01627359725534916
step: 140, loss: 0.003586382372304797
step: 150, loss: 0.10857612639665604
step: 160, loss: 0.010573367588222027
step: 170, loss: 0.038761306554079056
step: 180, loss: 0.00392323499545455
step: 190, loss: 0.023202674463391304
step: 200, loss: 0.0119552006945014
step: 210, loss: 0.010141120292246342
/usr/local/lib/python3.7/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=0.6817325800376649, f1=0.6802973977695167,
best_f1=0.7122736418511065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02947474643588066
step: 10, loss: 0.03233162313699722
step: 20, loss: 0.02629080042243004
step: 30, loss: 0.07962393760681152
step: 40, loss: 0.014145899564027786
step: 50, loss: 0.0033642041962593794
step: 60, loss: 0.013856588862836361
step: 70, loss: 0.0018395730294287205
step: 80, loss: 0.01776384562253952
step: 90, loss: 0.18475812673568726
step: 100, loss: 0.006199002265930176
step: 110, loss: 0.033172331750392914
step: 120, loss: 0.04217151924967766
step: 130, loss: 0.0002854851190932095
step: 140, loss: 0.04718365892767906

```

```

step: 150, loss: 0.0461152046918869
step: 160, loss: 0.00012994381540920585
step: 170, loss: 0.025985492393374443
step: 180, loss: 0.0011671045795083046
step: 190, loss: 0.03862963244318962
step: 200, loss: 0.00010398722224636003
step: 210, loss: 0.0029282760806381702
/usr/local/lib/python3.7/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=0.6765327695560254, f1=0.7016806722689076,
best_f1=0.7122736418511065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001841725897975266
step: 10, loss: 0.008140085265040398
step: 20, loss: 0.010334069840610027
step: 30, loss: 0.00018107492360286415
step: 40, loss: 8.361219806829467e-05
step: 50, loss: 0.005835505202412605
step: 60, loss: 0.00024216232122853398
step: 70, loss: 0.005031818524003029
step: 80, loss: 0.0797152891755104
step: 90, loss: 0.020999012514948845
step: 100, loss: 0.007354418747127056
step: 110, loss: 0.02921232394874096
step: 120, loss: 0.00017716911679599434
step: 130, loss: 0.0061330413445830345
step: 140, loss: 0.07918611913919449
step: 150, loss: 4.6285338612506166e-05
step: 160, loss: 0.003761367639526725
step: 170, loss: 0.08125654608011246
step: 180, loss: 0.019781583920121193
step: 190, loss: 0.043801210820674896
step: 200, loss: 0.002555664163082838
step: 210, loss: 0.00046464186743833125
/usr/local/lib/python3.7/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=0.6911447084233261, f1=0.7033898305084747,
best_f1=0.7122736418511065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009207571856677532
step: 10, loss: 0.007921477779746056
step: 20, loss: 0.00024121467140503228

```

```

step: 30, loss: 0.0018686670809984207
step: 40, loss: 0.014829765073955059
step: 50, loss: 0.26028409600257874
step: 60, loss: 0.0015618649777024984
step: 70, loss: 0.0002173230896005407
step: 80, loss: 0.012195532210171223
step: 90, loss: 0.0073429797776043415
step: 100, loss: 0.0033990852534770966
step: 110, loss: 0.01320671383291483
step: 120, loss: 0.020460141822695732
step: 130, loss: 0.0008394936448894441
step: 140, loss: 0.010939925909042358
step: 150, loss: 0.00014726321387570351
step: 160, loss: 0.042688049376010895
step: 170, loss: 0.015342648141086102
step: 180, loss: 0.0015214835293591022
step: 190, loss: 0.004441707860678434
step: 200, loss: 0.008466486819088459
step: 210, loss: 0.016885148361325264
/usr/local/lib/python3.7/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=0.7011952191235059, f1=0.694949494949495,
best_f1=0.7122736418511065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039151523378677666
step: 10, loss: 0.03428700193762779
step: 20, loss: 0.00018430595810059458
step: 30, loss: 0.00014328038378152996
step: 40, loss: 0.01459549367427826
step: 50, loss: 0.004470121581107378
step: 60, loss: 0.03668131306767464
step: 70, loss: 0.004468883853405714
step: 80, loss: 0.0004063703818246722
step: 90, loss: 0.0002059855469269678
step: 100, loss: 0.00022521983191836625
step: 110, loss: 0.00125539128202945
step: 120, loss: 6.54226605547592e-05
step: 130, loss: 0.00461843004450202
step: 140, loss: 0.005048183258622885
step: 150, loss: 0.006752829533070326
step: 160, loss: 0.0002217920555267483
step: 170, loss: 0.059409257024526596
step: 180, loss: 0.029297422617673874
step: 190, loss: 0.02966003492474556
step: 200, loss: 0.000881022890098393

```

```

step: 210, loss: 0.024308592081069946
/usr/local/lib/python3.7/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=0.6936170212765959, f1=0.6918238993710691,
best_f1=0.7122736418511065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00273239822126925
step: 10, loss: 0.00013430078979581594
step: 20, loss: 0.09146498888731003
step: 30, loss: 0.014344319701194763
step: 40, loss: 0.015073941089212894
step: 50, loss: 0.0011391581501811743
step: 60, loss: 0.02762358821928501
step: 70, loss: 0.00082786800339818
step: 80, loss: 0.060034994035959244
step: 90, loss: 0.00019072885334026068
step: 100, loss: 0.00029249401995912194
step: 110, loss: 0.018920741975307465
step: 120, loss: 0.0018334845080971718
step: 130, loss: 0.01182513777166605
step: 140, loss: 0.0026147174648940563
step: 150, loss: 0.007374620996415615
step: 160, loss: 0.0006962380139157176
step: 170, loss: 0.03256535902619362
step: 180, loss: 0.0002769119164440781
step: 190, loss: 0.003042082069441676
step: 200, loss: 0.001085385913029313
step: 210, loss: 0.019631389528512955
/usr/local/lib/python3.7/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=0.6852248394004283, f1=0.6877637130801688,
best_f1=0.7122736418511065
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:09, 234.25it/s]
load_f1 = 0.7186858316221767
real_f1 = 0.7128309572301426
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 133.36it/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 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: 0.5007076859474182


```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4459547698497772
step: 20, loss: 0.2530844807624817
step: 30, loss: 0.4085409343242645
step: 40, loss: 0.24701668322086334
step: 50, loss: 0.3180234432220459
step: 60, loss: 0.44196611642837524
step: 70, loss: 0.42545729875564575
step: 80, loss: 0.1960190385580063
step: 90, loss: 0.33640381693840027
step: 100, loss: 0.46481791138648987
step: 110, loss: 0.2363588660955429
step: 120, loss: 0.34485432505607605
step: 130, loss: 0.3104027211666107
step: 140, loss: 0.16757945716381073
step: 150, loss: 0.3541579246520996
step: 160, loss: 0.23288331925868988
step: 170, loss: 0.3918638825416565
step: 180, loss: 0.16666948795318604
step: 190, loss: 0.16701985895633698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.17255252570406795, f1=0.17048003589053387,
best_f1=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40477922558784485
step: 10, loss: 0.32801973819732666
step: 20, loss: 0.6068404912948608
step: 30, loss: 0.2753237783908844
step: 40, loss: 0.5731513500213623
step: 50, loss: 0.3156569302082062
step: 60, loss: 0.4718420207500458
step: 70, loss: 0.3161636292934418
step: 80, loss: 0.1690746545791626
step: 90, loss: 0.3215494155883789
step: 100, loss: 0.2501915693283081
step: 110, loss: 0.38812604546546936
step: 120, loss: 0.23445454239845276
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5171297788619995
step: 140, loss: 0.3405808210372925

```

```

step: 150, loss: 0.31836560368537903
step: 160, loss: 0.30105650424957275
step: 170, loss: 0.23355433344841003
step: 180, loss: 0.1628669798374176
step: 190, loss: 0.23815403878688812
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3972727060317993
step: 10, loss: 0.3828044533729553
step: 20, loss: 0.47094929218292236
step: 30, loss: 0.3190361559391022
step: 40, loss: 0.08152036368846893
step: 50, loss: 0.3771836757659912
step: 60, loss: 0.15875525772571564
step: 70, loss: 0.3719334602355957
step: 80, loss: 0.32725751399993896
step: 90, loss: 0.3997325301170349
step: 100, loss: 0.5375137329101562
step: 110, loss: 0.658149778842926
step: 120, loss: 0.3777647912502289
step: 130, loss: 0.15119940042495728
step: 140, loss: 0.36919766664505005
step: 150, loss: 0.3185155987739563
step: 160, loss: 0.6370458006858826
step: 170, loss: 0.43450748920440674
step: 180, loss: 0.36953723430633545
step: 190, loss: 0.16810502111911774
/usr/local/lib/python3.7/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=0.17239839213934793, f1=0.1748980516538287,
best_f1=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23526321351528168
step: 10, loss: 0.23206181824207306
step: 20, loss: 0.27644479274749756
step: 30, loss: 0.23028959333896637
step: 40, loss: 0.4967408776283264
step: 50, loss: 0.2361537516117096
step: 60, loss: 0.3480335474014282

```

```

step: 70, loss: 0.2837323546409607
step: 80, loss: 0.18955522775650024
step: 90, loss: 0.2447822242975235
step: 100, loss: 0.31348609924316406
step: 110, loss: 0.3669421076774597
step: 120, loss: 0.2621801197528839
step: 130, loss: 0.4680860638618469
step: 140, loss: 0.30352982878685
step: 150, loss: 0.23817887902259827
step: 160, loss: 0.3075494170188904
step: 170, loss: 0.41785845160484314
step: 180, loss: 0.3694458305835724
step: 190, loss: 0.16925355792045593
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40064534544944763
step: 10, loss: 0.3883371651172638
step: 20, loss: 0.17270240187644958
step: 30, loss: 0.11197365075349808
step: 40, loss: 0.31807178258895874
step: 50, loss: 0.5032292604446411
step: 60, loss: 0.232312873005867
step: 70, loss: 0.381689190864563
step: 80, loss: 0.37367257475852966
step: 90, loss: 0.32361888885498047
step: 100, loss: 0.4623410999774933
step: 110, loss: 0.4052218198776245
step: 120, loss: 0.23242688179016113
step: 130, loss: 0.5517232418060303
step: 140, loss: 0.3754287362098694
step: 150, loss: 0.30560022592544556
step: 160, loss: 0.16417355835437775
step: 170, loss: 0.3887691795825958
step: 180, loss: 0.23714114725589752
step: 190, loss: 0.31041255593299866
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3052140772342682
step: 10, loss: 0.2369353175163269
step: 20, loss: 0.31058356165885925
step: 30, loss: 0.4857231080532074
step: 40, loss: 0.24985039234161377
step: 50, loss: 0.31661996245384216
step: 60, loss: 0.44560956954956055
step: 70, loss: 0.3135412037372589
step: 80, loss: 0.32377052307128906
step: 90, loss: 0.2372400462627411
step: 100, loss: 0.48824888467788696
step: 110, loss: 0.23683682084083557
step: 120, loss: 0.47564244270324707
step: 130, loss: 0.5283013582229614
step: 140, loss: 0.18580381572246552
step: 150, loss: 0.37982189655303955
step: 160, loss: 0.380164235830307
step: 170, loss: 0.37975582480430603
step: 180, loss: 0.17126165330410004
step: 190, loss: 0.3109525442123413
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3086976110935211
step: 10, loss: 0.3060015141963959
step: 20, loss: 0.3109346926212311
step: 30, loss: 0.17587150633335114
step: 40, loss: 0.30725908279418945
step: 50, loss: 0.09185096621513367
step: 60, loss: 0.16362959146499634
step: 70, loss: 0.1575886309146881
step: 80, loss: 0.23599207401275635
step: 90, loss: 0.2293141782283783
step: 100, loss: 0.5574200749397278
step: 110, loss: 0.4442408084869385
step: 120, loss: 0.38333991169929504
step: 130, loss: 0.3137610852718353
step: 140, loss: 0.24428768455982208
step: 150, loss: 0.3082848787307739
step: 160, loss: 0.3854793310165405
step: 170, loss: 0.31059181690216064
step: 180, loss: 0.24842101335525513

```

```

step: 190, loss: 0.3182533085346222
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24006704986095428
step: 10, loss: 0.3781687319278717
step: 20, loss: 0.2326825112104416
step: 30, loss: 0.38109296560287476
step: 40, loss: 0.1723642200231552
step: 50, loss: 0.3095046579837799
step: 60, loss: 0.4678631126880646
step: 70, loss: 0.22987988591194153
step: 80, loss: 0.3806397020816803
step: 90, loss: 0.23732726275920868
step: 100, loss: 0.321940541267395
step: 110, loss: 0.38267597556114197
step: 120, loss: 0.38259634375572205
step: 130, loss: 0.30757880210876465
step: 140, loss: 0.38366034626960754
step: 150, loss: 0.32096388936042786
step: 160, loss: 0.18934640288352966
step: 170, loss: 0.2454599142074585
step: 180, loss: 0.3176250457763672
step: 190, loss: 0.308189332485199
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25313830375671387
step: 10, loss: 0.09670395404100418
step: 20, loss: 0.31908291578292847
step: 30, loss: 0.18397752940654755
step: 40, loss: 0.31320375204086304
step: 50, loss: 0.3816562592983246
step: 60, loss: 0.3762635886669159
step: 70, loss: 0.15730902552604675
step: 80, loss: 0.39060845971107483
step: 90, loss: 0.6956614255905151
step: 100, loss: 0.37872523069381714

```

```

step: 110, loss: 0.37021341919898987
step: 120, loss: 0.6058959364891052
step: 130, loss: 0.3038887083530426
step: 140, loss: 0.3131018877029419
step: 150, loss: 0.2440643608570099
step: 160, loss: 0.23825031518936157
step: 170, loss: 0.4530356228351593
step: 180, loss: 0.31804195046424866
step: 190, loss: 0.23979917168617249
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23567375540733337
step: 10, loss: 0.16619610786437988
step: 20, loss: 0.31651344895362854
step: 30, loss: 0.38185322284698486
step: 40, loss: 0.316760390996933
step: 50, loss: 0.09764513373374939
step: 60, loss: 0.3152846097946167
step: 70, loss: 0.3103924095630646
step: 80, loss: 0.3755863308906555
step: 90, loss: 0.16274692118167877
step: 100, loss: 0.2315405309200287
step: 110, loss: 0.23982249200344086
step: 120, loss: 0.38075515627861023
step: 130, loss: 0.505870521068573
step: 140, loss: 0.3810068666934967
step: 150, loss: 0.3158346712589264
step: 160, loss: 0.2378702163696289
step: 170, loss: 0.38198143243789673
step: 180, loss: 0.2360994815826416
step: 190, loss: 0.1626291573047638
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5308602452278137
step: 10, loss: 0.09615900367498398
step: 20, loss: 0.23873034119606018

```

```

step: 30, loss: 0.2409273386001587
step: 40, loss: 0.08992338180541992
step: 50, loss: 0.30952420830726624
step: 60, loss: 0.2344598025083542
step: 70, loss: 0.6248461008071899
step: 80, loss: 0.38936126232147217
step: 90, loss: 0.31795576214790344
step: 100, loss: 0.1673242747783661
step: 110, loss: 0.24601615965366364
step: 120, loss: 0.3158673942089081
step: 130, loss: 0.6071608662605286
step: 140, loss: 0.4625231623649597
step: 150, loss: 0.31180381774902344
step: 160, loss: 0.3096649646759033
step: 170, loss: 0.43518343567848206
step: 180, loss: 0.3149806559085846
step: 190, loss: 0.10613434761762619
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3945658504962921
step: 10, loss: 0.31505754590034485
step: 20, loss: 0.5459966063499451
step: 30, loss: 0.374648779630661
step: 40, loss: 0.31627053022384644
step: 50, loss: 0.5318130850791931
step: 60, loss: 0.45587825775146484
step: 70, loss: 0.37685278058052063
step: 80, loss: 0.3832031488418579
step: 90, loss: 0.4486667811870575
step: 100, loss: 0.30504322052001953
step: 110, loss: 0.452065110206604
step: 120, loss: 0.17171962559223175
step: 130, loss: 0.24183616042137146
step: 140, loss: 0.38008707761764526
step: 150, loss: 0.3813709616661072
step: 160, loss: 0.24158687889575958
step: 170, loss: 0.4555191993713379
step: 180, loss: 0.31267842650413513
step: 190, loss: 0.16776160895824432
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3075992465019226
step: 10, loss: 0.31165367364883423
step: 20, loss: 0.3978871703147888
step: 30, loss: 0.38375744223594666
step: 40, loss: 0.252311110496521
step: 50, loss: 0.24602334201335907
step: 60, loss: 0.3045025169849396
step: 70, loss: 0.16587567329406738
step: 80, loss: 0.307273805141449
step: 90, loss: 0.3094937205314636
step: 100, loss: 0.1633903533220291
step: 110, loss: 0.6149309277534485
step: 120, loss: 0.1641232818365097
step: 130, loss: 0.24030044674873352
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.3103846609592438
step: 150, loss: 0.38085001707077026
step: 160, loss: 0.45359018445014954
step: 170, loss: 0.1004386618733406
step: 180, loss: 0.2537080943584442
step: 190, loss: 0.31165584921836853
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3128262758255005
step: 10, loss: 0.2400529831647873
step: 20, loss: 0.3783864378929138
step: 30, loss: 0.103425994515419
step: 40, loss: 0.31919822096824646
step: 50, loss: 0.3754116892814636
step: 60, loss: 0.31504368782043457
step: 70, loss: 0.23842044174671173
step: 80, loss: 0.3837747573852539
step: 90, loss: 0.16405175626277924
step: 100, loss: 0.30378931760787964
step: 110, loss: 0.4500805139541626
step: 120, loss: 0.3148129880428314
step: 130, loss: 0.3179306983947754

```



```

step: 140, loss: 0.16698607802391052
step: 150, loss: 0.4426206946372986
step: 160, loss: 0.384696364402771
step: 170, loss: 0.24742300808429718
step: 180, loss: 0.3203296363353729
step: 190, loss: 0.5346062779426575
/usr/local/lib/python3.7/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=0.17048003589053387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5377542972564697
step: 10, loss: 0.23658937215805054
step: 20, loss: 0.23953913152217865
step: 30, loss: 0.3160841166973114
step: 40, loss: 0.6619205474853516
step: 50, loss: 0.3124125897884369
step: 60, loss: 0.1799473911523819
step: 70, loss: 0.24195724725723267
step: 80, loss: 0.24034211039543152
step: 90, loss: 0.3096872866153717
step: 100, loss: 0.38869622349739075
step: 110, loss: 0.4583221673965454
step: 120, loss: 0.23766478896141052
step: 130, loss: 0.3093828558921814
step: 140, loss: 0.4482371211051941
step: 150, loss: 0.1678023487329483
step: 160, loss: 0.31115344166755676
step: 170, loss: 0.7470533847808838
step: 180, loss: 0.306886225938797
step: 190, loss: 0.4529626965522766
/usr/local/lib/python3.7/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=0.17048003589053387
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:13, 150.78it/s]
load_f1 = 0.17439192290041303
real_f1 = 0.17255252570406795
733it [00:00, 3598.05it/s]/usr/local/lib/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, 133.03it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

```

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

```

```

--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

```
step: 0, loss: 0.4932442605495453
```

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

```
step: 10, loss: 0.43414077162742615
```

```
step: 20, loss: 0.3211682140827179
```

```
step: 30, loss: 0.3994652330875397
```

```
step: 40, loss: 0.5873842239379883
```

```
step: 50, loss: 0.3814600706100464
```

```
step: 60, loss: 0.5959380269050598
```

```
step: 70, loss: 0.29110825061798096
```

```
step: 80, loss: 0.2785693109035492
```

```
step: 90, loss: 0.22751769423484802
```

```
step: 100, loss: 0.18110810220241547
```

```
step: 110, loss: 0.4132592976093292
```

```
step: 120, loss: 0.3223961889743805
```

```
step: 130, loss: 0.3354262709617615
```

```
step: 140, loss: 0.42585447430610657
```

```
step: 150, loss: 0.32193437218666077
```

```
step: 160, loss: 0.3678007125854492
```

```
step: 170, loss: 0.28990066051483154
```

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

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

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

```
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: 0.30684348940849304
```

```
step: 10, loss: 0.4049180746078491
```

```
step: 20, loss: 0.27220144867897034
```

```
step: 30, loss: 0.30201393365859985
```

```
step: 40, loss: 0.0699823647737503
```

```
step: 50, loss: 0.42761796712875366
```

```
step: 60, loss: 0.186068594455719
```

```
step: 70, loss: 0.4902538061141968
```

```
step: 80, loss: 0.21498553454875946
```

```
step: 90, loss: 0.24257788062095642
```

```
step: 100, loss: 0.5041926503181458
```

```
step: 110, loss: 0.28197693824768066
```

```
step: 120, loss: 0.21297800540924072
```

```

step: 130, loss: 0.386868953704834
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.5315848588943481
step: 150, loss: 0.4360721707344055
step: 160, loss: 0.41331958770751953
step: 170, loss: 0.2959328293800354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.39516129032258074, f1=0.40909090909090906,
best_f1=0.40909090909090906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45995423197746277
step: 10, loss: 0.30798113346099854
step: 20, loss: 0.23314902186393738
step: 30, loss: 0.25919392704963684
step: 40, loss: 0.3886798322200775
step: 50, loss: 0.5570719242095947
step: 60, loss: 0.3106415867805481
step: 70, loss: 0.2314510941505432
step: 80, loss: 0.3556460440158844
step: 90, loss: 0.44598668813705444
step: 100, loss: 0.31536272168159485
step: 110, loss: 0.2139481008052826
step: 120, loss: 0.560570240020752
step: 130, loss: 0.5038101673126221
step: 140, loss: 0.4356042146682739
step: 150, loss: 0.18908172845840454
step: 160, loss: 0.16809570789337158
step: 170, loss: 0.2626219093799591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.709016393442623, f1=0.6958333333333333,
best_f1=0.6958333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.11472737044095993
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 10, loss: 0.23041003942489624
step: 20, loss: 0.22430765628814697
step: 30, loss: 0.3949785828590393
step: 40, loss: 0.211504265666008
step: 50, loss: 0.30994912981987
step: 60, loss: 0.42967233061790466
step: 70, loss: 0.0919245183467865
step: 80, loss: 0.24325062334537506
step: 90, loss: 0.22048918902873993
step: 100, loss: 0.3313126564025879
step: 110, loss: 0.2771454155445099
step: 120, loss: 0.34647345542907715
step: 130, loss: 0.08287953585386276
step: 140, loss: 0.07364823669195175
step: 150, loss: 0.24026569724082947
step: 160, loss: 0.010898357257246971
step: 170, loss: 0.14613254368305206
/usr/local/lib/python3.7/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=0.700228832951945, f1=0.6917960088691796,
best_f1=0.6958333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13686133921146393
step: 10, loss: 0.18151749670505524
step: 20, loss: 0.1246914342045784
step: 30, loss: 0.036801353096961975
step: 40, loss: 0.13547809422016144
step: 50, loss: 0.08449704945087433
step: 60, loss: 0.11022602766752243
step: 70, loss: 0.04253828153014183
step: 80, loss: 0.003641596296802163
step: 90, loss: 0.11466703563928604
step: 100, loss: 0.11842340975999832
step: 110, loss: 0.2168886661529541
step: 120, loss: 0.12287845462560654
step: 130, loss: 0.05718589946627617
step: 140, loss: 0.09755487740039825
step: 150, loss: 0.06737519800662994
step: 160, loss: 0.22263318300247192
step: 170, loss: 0.07067116349935532
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7898734177215191, f1=0.8175182481751825,
best_f1=0.8175182481751825
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22376327216625214
step: 10, loss: 0.1540909856557846
step: 20, loss: 0.14980334043502808
step: 30, loss: 0.039026446640491486
step: 40, loss: 0.0567476786673069
step: 50, loss: 0.014475611038506031
step: 60, loss: 0.05112658813595772
step: 70, loss: 0.03319849073886871
step: 80, loss: 0.034398067742586136
step: 90, loss: 0.06262776255607605
step: 100, loss: 0.008532890118658543
step: 110, loss: 0.13058525323867798
step: 120, loss: 0.09125262498855591
step: 130, loss: 0.03352919965982437
step: 140, loss: 0.06815748661756516
step: 150, loss: 0.025641564279794693
step: 160, loss: 0.042408064007759094
step: 170, loss: 0.07322967797517776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8410256410256409, f1=0.8592592592592592,
best_f1=0.8592592592592592
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017558416351675987
step: 10, loss: 0.013604589737951756
step: 20, loss: 0.008813215419650078
step: 30, loss: 0.005815763026475906
step: 40, loss: 0.0017528635216876864
step: 50, loss: 0.006246916949748993
step: 60, loss: 0.06472475081682205
step: 70, loss: 0.010304760187864304
step: 80, loss: 0.039191197603940964

```

```

step: 90, loss: 0.08585865795612335
step: 100, loss: 0.0021184636279940605
step: 110, loss: 0.1560000628232956
step: 120, loss: 0.029381858184933662
step: 130, loss: 0.01010277308523655
step: 140, loss: 0.04906206950545311
step: 150, loss: 0.17576493322849274
step: 160, loss: 0.021084817126393318
step: 170, loss: 0.03455565124750137
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8493827160493828, f1=0.8578431372549019,
best_f1=0.8578431372549019
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10514428466558456
step: 10, loss: 0.020683633163571358
step: 20, loss: 0.011805866844952106
step: 30, loss: 0.002316382946446538
step: 40, loss: 0.053487591445446014
step: 50, loss: 0.004890670068562031
step: 60, loss: 0.021500730887055397
step: 70, loss: 0.11722476035356522
step: 80, loss: 0.0029200599528849125
step: 90, loss: 0.13122588396072388
step: 100, loss: 0.006951712071895599
step: 110, loss: 0.09250941872596741
step: 120, loss: 0.31853818893432617
step: 130, loss: 0.04395592212677002
step: 140, loss: 0.04478892683982849
step: 150, loss: 0.0027022778522223234
step: 160, loss: 0.004210892133414745
step: 170, loss: 0.19539359211921692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8592964824120604, f1=0.8704156479217603,
best_f1=0.8704156479217603

```


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

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

step: 0, loss: 0.03312435373663902
step: 10, loss: 0.006720793899148703
step: 20, loss: 0.008728583343327045
step: 30, loss: 0.0036116153933107853
step: 40, loss: 0.0163229089230299
step: 50, loss: 0.00787121057510376
step: 60, loss: 0.04084034636616707
step: 70, loss: 0.12181013077497482
step: 80, loss: 0.017401494085788727
step: 90, loss: 0.10766280442476273
step: 100, loss: 0.04677635058760643
step: 110, loss: 0.028482364490628242
step: 120, loss: 0.0029274169355630875
step: 130, loss: 0.004758235067129135
step: 140, loss: 0.04156806319952011
step: 150, loss: 0.20777671039104462
step: 160, loss: 0.00398056348785758
step: 170, loss: 0.0796554833650589

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

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

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.

warnings.warn(SAVE_STATE_WARNING, UserWarning)

epoch 9: dev_f1=0.8678304239401496, f1=0.8687350835322196,
best_f1=0.8687350835322196

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

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

step: 0, loss: 0.0011619583237916231
step: 10, loss: 0.0033235023729503155
step: 20, loss: 0.0900583490729332
step: 30, loss: 0.0702228918671608
step: 40, loss: 0.014164626598358154
step: 50, loss: 0.06059402972459793
step: 60, loss: 0.08149894326925278
step: 70, loss: 0.0029767516534775496
step: 80, loss: 0.05278920382261276
step: 90, loss: 0.0011063176207244396
step: 100, loss: 0.0012385299196466804
step: 110, loss: 0.013927144929766655
step: 120, loss: 0.025178125128149986
step: 130, loss: 0.013607967644929886
step: 140, loss: 0.0004114362527616322

```

step: 150, loss: 0.15678073465824127
step: 160, loss: 0.0004804708587471396
step: 170, loss: 0.06220114603638649
/usr/local/lib/python3.7/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=0.834567901234568, f1=0.8613138686131387,
best_f1=0.8687350835322196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0328134149312973
step: 10, loss: 0.001213441020809114
step: 20, loss: 0.008714229799807072
step: 30, loss: 0.029717596247792244
step: 40, loss: 0.00045711532584391534
step: 50, loss: 0.0693456381559372
step: 60, loss: 0.01845865324139595
step: 70, loss: 0.0019800099544227123
step: 80, loss: 0.019895091652870178
step: 90, loss: 0.0007769836811348796
step: 100, loss: 0.07291798293590546
step: 110, loss: 0.004318976309150457
step: 120, loss: 0.013558941893279552
step: 130, loss: 0.007137123495340347
step: 140, loss: 0.03048340231180191
step: 150, loss: 0.0005200274754315615
step: 160, loss: 0.026152899488806725
step: 170, loss: 0.004063330590724945
/usr/local/lib/python3.7/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=0.860759493670886, f1=0.880195599022005,
best_f1=0.8687350835322196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006293224869295955
step: 10, loss: 0.0005138298729434609
step: 20, loss: 0.0005732111749239266
step: 30, loss: 0.10968022793531418
step: 40, loss: 0.001476326840929687
step: 50, loss: 0.0010075642494484782
step: 60, loss: 0.0009901663288474083
step: 70, loss: 0.00024872113135643303
step: 80, loss: 0.0005156948463991284
step: 90, loss: 0.03606536611914635
step: 100, loss: 0.0012282003881409764

```

```

step: 110, loss: 0.02517501451075077
step: 120, loss: 0.01431957632303238
step: 130, loss: 0.06052220240235329
step: 140, loss: 0.0010125101543962955
step: 150, loss: 0.0018945835763588548
step: 160, loss: 0.07867544144392014
step: 170, loss: 0.0011339327320456505
/usr/local/lib/python3.7/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=0.858611825192802, f1=0.875912408759124,
best_f1=0.8687350835322196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002837914798874408
step: 10, loss: 0.0002376153861405328
step: 20, loss: 0.00023176425020210445
step: 30, loss: 0.03101784735918045
step: 40, loss: 0.018246037885546684
step: 50, loss: 0.030677849426865578
step: 60, loss: 0.004618777893483639
step: 70, loss: 0.08210950344800949
step: 80, loss: 0.0003739110834430903
step: 90, loss: 0.001263173995539546
step: 100, loss: 0.021733099594712257
step: 110, loss: 0.005146802868694067
step: 120, loss: 0.018439000472426414
step: 130, loss: 0.0002832514001056552
step: 140, loss: 0.04594472795724869
step: 150, loss: 0.0010528960265219212
step: 160, loss: 0.0007387652294710279
step: 170, loss: 0.0007882682839408517
/usr/local/lib/python3.7/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=0.854922279792746, f1=0.8671679197994987,
best_f1=0.8687350835322196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0046487911604344845
step: 10, loss: 0.00046072682016529143
step: 20, loss: 0.02359727956354618
step: 30, loss: 0.0027908135671168566
step: 40, loss: 0.0007842873455956578
step: 50, loss: 0.0006531162071041763
step: 60, loss: 0.028054602444171906

```

```

step: 70, loss: 0.005899031646549702
step: 80, loss: 0.004108994267880917
step: 90, loss: 0.00028283221763558686
step: 100, loss: 0.013401279225945473
step: 110, loss: 0.02063119225203991
step: 120, loss: 0.002292754827067256
step: 130, loss: 0.041052959859371185
step: 140, loss: 0.02993522211909294
step: 150, loss: 0.026346122846007347
step: 160, loss: 0.0002600476727820933
step: 170, loss: 0.06786530464887619
/usr/local/lib/python3.7/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=0.8459530026109661, f1=0.8614609571788413,
best_f1=0.8687350835322196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019956508185714483
step: 10, loss: 0.0018712085438892245
step: 20, loss: 0.0002519749687053263
step: 30, loss: 0.010671284981071949
step: 40, loss: 0.0003977244778070599
step: 50, loss: 0.001885620760731399
step: 60, loss: 0.00550888804718852
step: 70, loss: 0.06777983158826828
step: 80, loss: 0.0008853843901306391
step: 90, loss: 0.001251407666131854
step: 100, loss: 0.040335461497306824
step: 110, loss: 0.0042247651144862175
step: 120, loss: 0.0017865129048004746
step: 130, loss: 0.004164399579167366
step: 140, loss: 0.014545539394021034
step: 150, loss: 0.0006157098687253892
step: 160, loss: 0.0002392939495621249
step: 170, loss: 0.0006241678493097425
/usr/local/lib/python3.7/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=0.8556701030927835, f1=0.8663366336633663,
best_f1=0.8687350835322196
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 [00:06, 297.07it/s]
load_f1 = 0.865
real_f1 = 0.8656716417910448
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 [00:23, 189.38it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

```

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

```

```

--finetuning \
--lm roberta \
--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.5878954529762268

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

step: 10, loss: 0.41241997480392456

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5145020484924316
step: 30, loss: 0.37285923957824707
step: 40, loss: 0.2193586677312851
step: 50, loss: 0.12826663255691528
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.15190435945987701
step: 70, loss: 0.07786142081022263
step: 80, loss: 0.12642160058021545
step: 90, loss: 0.1976889818906784
step: 100, loss: 0.0834370106458664
step: 110, loss: 0.10277233272790909
step: 120, loss: 0.07762753218412399
step: 130, loss: 0.032531850039958954
step: 140, loss: 0.02484065480530262
step: 150, loss: 0.1321011334657669
step: 160, loss: 0.01798819750547409
step: 170, loss: 0.14284339547157288
step: 180, loss: 0.08303587883710861
step: 190, loss: 0.035045892000198364
step: 200, loss: 0.02857382223010063
step: 210, loss: 0.044332969933748245
step: 220, loss: 0.07466381788253784
step: 230, loss: 0.009041416458785534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9547511312217195, 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: 0.01962389051914215
step: 10, loss: 0.12038733810186386
step: 20, loss: 0.03418748453259468
step: 30, loss: 0.03367840498685837
step: 40, loss: 0.07524746656417847
step: 50, loss: 0.00409417599439621
step: 60, loss: 0.004430041648447514
step: 70, loss: 0.015764933079481125
step: 80, loss: 0.028223449364304543
step: 90, loss: 0.04968501254916191
step: 100, loss: 0.0834604948759079
step: 110, loss: 0.04136635363101959

```

```

step: 120, loss: 0.006389596499502659
step: 130, loss: 0.017808135598897934
step: 140, loss: 0.004884947557002306
step: 150, loss: 0.08347605168819427
step: 160, loss: 0.05089912936091423
step: 170, loss: 0.0031085682567209005
step: 180, loss: 0.018724653869867325
step: 190, loss: 0.0028734158258885145
step: 200, loss: 0.036374080926179886
step: 210, loss: 0.013099313713610172
step: 220, loss: 0.011424994096159935
step: 230, loss: 0.001351650571450591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9751131221719457, f1=0.9708520179372198,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010288966819643974
step: 10, loss: 0.001629338599741459
step: 20, loss: 0.0010919349733740091
step: 30, loss: 0.004635846242308617
step: 40, loss: 0.03393113613128662
step: 50, loss: 0.005609545856714249
step: 60, loss: 0.05713558569550514
step: 70, loss: 0.009027511812746525
step: 80, loss: 0.1385492980480194
step: 90, loss: 0.011628391221165657
step: 100, loss: 0.050059061497449875
step: 110, loss: 0.011689823120832443
step: 120, loss: 0.0006491562235169113
step: 130, loss: 0.00454377755522728
step: 140, loss: 0.0028504813089966774
step: 150, loss: 0.0693923681974411
step: 160, loss: 0.0034567771945148706
step: 170, loss: 0.0009243748500011861
step: 180, loss: 0.011067193001508713
step: 190, loss: 0.02692331001162529
step: 200, loss: 0.004339561332017183
step: 210, loss: 0.0026432329323142767
step: 220, loss: 0.11488202959299088
step: 230, loss: 0.02273985929787159

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9787234042553192, f1=0.9753363228699552,
best_f1=0.9753363228699552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006537520792335272
step: 10, loss: 0.005563098005950451
step: 20, loss: 0.0020574801601469517
step: 30, loss: 0.0017038540681824088
step: 40, loss: 0.09897862374782562
step: 50, loss: 0.02466108836233616
step: 60, loss: 0.004778763744980097
step: 70, loss: 0.005040252115577459
step: 80, loss: 0.1289103627204895
step: 90, loss: 0.02513430081307888
step: 100, loss: 0.01500006765127182
step: 110, loss: 0.0015422850847244263
step: 120, loss: 0.03330099955201149
step: 130, loss: 0.0020404020324349403
step: 140, loss: 0.001299368916079402
step: 150, loss: 0.004480555187910795
step: 160, loss: 0.0014153467491269112
step: 170, loss: 0.0006985919317230582
step: 180, loss: 0.16275212168693542
step: 190, loss: 0.0031694089993834496
step: 200, loss: 0.08149583637714386
step: 210, loss: 0.09919211268424988
step: 220, loss: 0.0026240318547934294
step: 230, loss: 0.002431543543934822
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.9797752808988766, f1=0.9785794813979707,
best_f1=0.9785794813979707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.026764731854200363
step: 10, loss: 0.01627005636692047
step: 20, loss: 0.0015794343780726194
step: 30, loss: 0.023849239572882652
step: 40, loss: 0.0013374063419178128
step: 50, loss: 0.0005781278596259654
step: 60, loss: 0.04711270332336426
step: 70, loss: 0.012696415185928345
step: 80, loss: 0.013963919132947922
step: 90, loss: 0.07354117184877396
step: 100, loss: 0.0006943641928955913
step: 110, loss: 0.02431516721844673
step: 120, loss: 0.001532247755676508
step: 130, loss: 0.005496657453477383
step: 140, loss: 0.003775998018682003
step: 150, loss: 0.017573732882738113
step: 160, loss: 0.0017541081178933382
step: 170, loss: 0.0034715288784354925
step: 180, loss: 0.003999467473477125
step: 190, loss: 0.03162749484181404
step: 200, loss: 0.0008532058564014733
step: 210, loss: 0.0015793911879882216
step: 220, loss: 0.0037518120370805264
step: 230, loss: 0.03939976170659065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, f1=0.9755011135857461,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000410274020396173
step: 10, loss: 0.0012413503136485815
step: 20, loss: 0.0024000133853405714
step: 30, loss: 0.00043938111048191786
step: 40, loss: 0.00015535415150225163
step: 50, loss: 0.022074764594435692
step: 60, loss: 0.0009381996933370829
step: 70, loss: 0.0018765977583825588
step: 80, loss: 0.0006077882135286927
step: 90, loss: 0.003266378305852413
step: 100, loss: 0.05624815821647644
step: 110, loss: 0.021918421611189842

```

```

step: 120, loss: 0.000759524351451546
step: 130, loss: 0.024275654926896095
step: 140, loss: 0.00027278336347080767
step: 150, loss: 0.005796312354505062
step: 160, loss: 0.005380483344197273
step: 170, loss: 0.0002824855619110167
step: 180, loss: 0.00020584408775903285
step: 190, loss: 0.0003990779514424503
step: 200, loss: 0.09050888568162918
step: 210, loss: 0.0005828112480230629
step: 220, loss: 0.007291209418326616
step: 230, loss: 0.022660205140709877
/usr/local/lib/python3.7/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=0.9741282339707535, f1=0.9742441209406495,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01698308438062668
step: 10, loss: 0.0005151015357114375
step: 20, loss: 0.0017106914892792702
step: 30, loss: 0.00042399606900289655
step: 40, loss: 0.0005609680665656924
step: 50, loss: 0.0037427989300340414
step: 60, loss: 0.0011351663852110505
step: 70, loss: 0.0004726093611679971
step: 80, loss: 0.00030487231560982764
step: 90, loss: 0.007244607899338007
step: 100, loss: 0.0006928913062438369
step: 110, loss: 0.0031361053697764874
step: 120, loss: 0.0007318432326428592
step: 130, loss: 0.002276051789522171
step: 140, loss: 0.00020786082313861698
step: 150, loss: 0.03276783600449562
step: 160, loss: 0.00020530739857349545
step: 170, loss: 0.0013928001280874014
step: 180, loss: 0.00013320757716428488
step: 190, loss: 0.012394984252750874
step: 200, loss: 0.00046376147656701505
step: 210, loss: 0.01646626926958561
step: 220, loss: 0.0004764232144225389
step: 230, loss: 0.0019399162847548723
/usr/local/lib/python3.7/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=0.9784824462061155, f1=0.9727891156462585,

```

```

best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004010174423456192
step: 10, loss: 0.0020230826921761036
step: 20, loss: 0.0011990227503702044
step: 30, loss: 0.0014596764231100678
step: 40, loss: 0.016731515526771545
step: 50, loss: 0.001512648887000978
step: 60, loss: 0.0007983383839018643
step: 70, loss: 0.00026268966030329466
step: 80, loss: 0.016244662925601006
step: 90, loss: 0.0006132968119345605
step: 100, loss: 0.000980628770776093
step: 110, loss: 0.00043996970634907484
step: 120, loss: 0.00034043638152070343
step: 130, loss: 0.0009713330073282123
step: 140, loss: 0.00041119990055449307
step: 150, loss: 0.04672855883836746
step: 160, loss: 0.0003158350591547787
step: 170, loss: 0.022235674783587456
step: 180, loss: 0.00014774040027987212
step: 190, loss: 0.00022174000332597643
step: 200, loss: 0.0009537247242406011
step: 210, loss: 0.0005705961957573891
step: 220, loss: 0.0002431282337056473
step: 230, loss: 0.0024392514023929834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853107344632768, f1=0.9808773903262092,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023157463874667883
step: 10, loss: 0.0005289289401844144
step: 20, loss: 0.0005725033697672188
step: 30, loss: 0.0004001847410108894
step: 40, loss: 0.00034286451409570873
step: 50, loss: 0.0004895210731774569
step: 60, loss: 0.0029470003210008144
step: 70, loss: 0.031788554042577744

```

```

step: 80, loss: 9.548783418722451e-05
step: 90, loss: 0.004599459934979677
step: 100, loss: 0.00018639561312738806
step: 110, loss: 0.00012981753388885409
step: 120, loss: 0.011225691065192223
step: 130, loss: 0.0020608180202543736
step: 140, loss: 0.00015349405293818563
step: 150, loss: 0.00011916097719222307
step: 160, loss: 0.0005434934282675385
step: 170, loss: 0.00028198337531648576
step: 180, loss: 0.00018788351735565811
step: 190, loss: 7.973817264428362e-05
step: 200, loss: 0.00013030362606514245
step: 210, loss: 0.0013321859296411276
step: 220, loss: 0.00037843550671823323
step: 230, loss: 0.007964476011693478
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9831649831649831, f1=0.9821428571428571,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019161620002705604
step: 10, loss: 0.00013083405792713165
step: 20, loss: 0.00025265978183597326
step: 30, loss: 7.433520659105852e-05
step: 40, loss: 0.0003244945255573839
step: 50, loss: 6.023654350428842e-05
step: 60, loss: 8.167113992385566e-05
step: 70, loss: 0.002007000381127
step: 80, loss: 0.00012335866631474346
step: 90, loss: 5.729562326450832e-05
step: 100, loss: 5.862085163244046e-05
step: 110, loss: 0.002173434942960739
step: 120, loss: 5.359021088224836e-05
step: 130, loss: 0.00010448430111864582
step: 140, loss: 7.172156620072201e-05
step: 150, loss: 0.0016942620277404785
step: 160, loss: 3.448602365097031e-05
step: 170, loss: 7.145263953134418e-05
step: 180, loss: 0.0001044680320774205
step: 190, loss: 6.687652785331011e-05
step: 200, loss: 8.943817374529317e-05
step: 210, loss: 0.0004661538696382195
step: 220, loss: 0.001986101036891341
step: 230, loss: 0.0002869142044801265

```

```

/usr/local/lib/python3.7/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=0.9821029082774049, f1=0.9755011135857461,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.438605174073018e-05
step: 10, loss: 0.00020683107140939683
step: 20, loss: 0.0002614754776004702
step: 30, loss: 0.00011008758883690462
step: 40, loss: 3.108824239461683e-05
step: 50, loss: 4.240229100105353e-05
step: 60, loss: 0.007983539253473282
step: 70, loss: 8.80743027664721e-05
step: 80, loss: 0.00933790858834982
step: 90, loss: 0.02897314354777336
step: 100, loss: 8.466361759928986e-05
step: 110, loss: 0.0038341055624186993
step: 120, loss: 6.87007704982534e-05
step: 130, loss: 5.278343451209366e-05
step: 140, loss: 0.007814290001988411
step: 150, loss: 7.068861305015162e-05
step: 160, loss: 0.04071531444787979
step: 170, loss: 0.0012350394390523434
step: 180, loss: 5.2749957831110805e-05
step: 190, loss: 4.199998511467129e-05
step: 200, loss: 0.0003901766613125801
step: 210, loss: 4.084520696778782e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 4.628994065569714e-05
step: 230, loss: 4.2988602217519656e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.983050847457627, f1=0.983050847457627,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.111614831956103e-05
step: 10, loss: 3.88523330911994e-05
step: 20, loss: 0.014386233873665333
step: 30, loss: 0.029033903032541275
step: 40, loss: 5.606487320619635e-05
step: 50, loss: 0.0005737087340094149
step: 60, loss: 0.000775650201831013

```

```

step: 70, loss: 3.689329241751693e-05
step: 80, loss: 2.690258224902209e-05
step: 90, loss: 0.012367328628897667
step: 100, loss: 5.4025702411308885e-05
step: 110, loss: 3.163330620736815e-05
step: 120, loss: 3.886761624016799e-05
step: 130, loss: 5.2774288633372635e-05
step: 140, loss: 4.03461599489674e-05
step: 150, loss: 4.3862593884114176e-05
step: 160, loss: 0.0003053421969525516
step: 170, loss: 0.00017686362843960524
step: 180, loss: 3.138039028272033e-05
step: 190, loss: 6.621033389819786e-05
step: 200, loss: 2.9219474527053535e-05
step: 210, loss: 0.00010676261445041746
step: 220, loss: 0.04606611654162407
step: 230, loss: 4.266224641469307e-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=0.9830124575311437, f1=0.9819004524886877,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.244561907602474e-05
step: 10, loss: 3.808881592703983e-05
step: 20, loss: 4.86450080643408e-05
step: 30, loss: 2.4790413590380922e-05
step: 40, loss: 3.9933027437655255e-05
step: 50, loss: 0.0008851459715515375
step: 60, loss: 5.3984171245247126e-05
step: 70, loss: 2.577422492322512e-05
step: 80, loss: 2.4179480533348396e-05
step: 90, loss: 5.1650502427946776e-05
step: 100, loss: 0.00010522431693971157
step: 110, loss: 0.00031390131334774196
step: 120, loss: 3.4595221222843975e-05
step: 130, loss: 3.611122883739881e-05
step: 140, loss: 4.334430195740424e-05
step: 150, loss: 1.7444872355554253e-05
step: 160, loss: 0.010167445987462997
step: 170, loss: 3.7109482946107164e-05
step: 180, loss: 0.036417316645383835
step: 190, loss: 5.19633867952507e-05
step: 200, loss: 1.5761374015710317e-05
step: 210, loss: 0.0012578062014654279
step: 220, loss: 3.133958671241999e-05

```

```

step: 230, loss: 2.9447204724419862e-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=0.9853438556933484, f1=0.9807909604519773,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.178588303853758e-05
step: 10, loss: 2.3882012101239525e-05
step: 20, loss: 8.042727131396532e-05
step: 30, loss: 4.498018097365275e-05
step: 40, loss: 2.7662796128424816e-05
step: 50, loss: 2.7800573661806993e-05
step: 60, loss: 2.8936894523212686e-05
step: 70, loss: 2.8709444450214505e-05
step: 80, loss: 3.051126259379089e-05
step: 90, loss: 3.8018362829461694e-05
step: 100, loss: 3.0229304684326053e-05
step: 110, loss: 3.8178728573257104e-05
step: 120, loss: 1.5265952242771164e-05
step: 130, loss: 5.2667423005914316e-05
step: 140, loss: 3.586460661608726e-05
step: 150, loss: 1.9505061572999693e-05
step: 160, loss: 4.878581239609048e-05
step: 170, loss: 4.135460403631441e-05
step: 180, loss: 2.9752713089692406e-05
step: 190, loss: 2.134904934791848e-05
step: 200, loss: 3.579078475013375e-05
step: 210, loss: 2.5025570721481927e-05
step: 220, loss: 3.5798024327959865e-05
step: 230, loss: 0.007005632389336824
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9819413092550789,
best_f1=0.9807909604519773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019370444118976593
step: 10, loss: 3.461351298028603e-05
step: 20, loss: 7.556763739557937e-05

```



```

step: 30, loss: 3.234874384361319e-05
step: 40, loss: 2.1516605556826107e-05
step: 50, loss: 2.8054106223862618e-05
step: 60, loss: 0.026017872616648674
step: 70, loss: 2.5241723051294684e-05
step: 80, loss: 3.9194284909171984e-05
step: 90, loss: 3.052696047234349e-05
step: 100, loss: 2.2283878934103996e-05
step: 110, loss: 3.0885224987287074e-05
step: 120, loss: 0.034610822796821594
step: 130, loss: 2.4511538867955096e-05
step: 140, loss: 0.0063927024602890015
step: 150, loss: 0.00040201330557465553
step: 160, loss: 0.007931732572615147
step: 170, loss: 2.0115947336307727e-05
step: 180, loss: 2.8769321943400428e-05
step: 190, loss: 2.7535283152246848e-05
step: 200, loss: 5.807418347103521e-05
step: 210, loss: 0.01512727327644825
step: 220, loss: 2.671311995072756e-05
step: 230, loss: 3.669166471809149e-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=0.9842342342342343, f1=0.9808773903262092,
best_f1=0.9807909604519773
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 188.86it/s]
load_f1 = 0.9842342342342343
real_f1 = 0.9831271091113611
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:24, 176.28it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

Downloading: 100% 481/481 [00:00<00:00, 413kB/s]
 Downloading: 100% 899k/899k [00:01<00:00, 791kB/s]
 Downloading: 100% 456k/456k [00:00<00:00, 469kB/s]
 Downloading: 100% 501M/501M [00:07<00:00, 69.6MB/s]
 Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

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

After processing overrides, optimization options are:

```

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

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

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

step: 0, loss: 0.6306542754173279

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: 0.4551503658294678

step: 20, loss: 0.2890413701534271

step: 30, loss: 0.354692667722702

step: 40, loss: 0.39851394295692444

step: 50, loss: 0.5084888339042664

step: 60, loss: 0.2513446807861328

step: 70, loss: 0.2024441659450531

step: 80, loss: 0.16071867942810059

step: 90, loss: 0.1744507998228073

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

step: 100, loss: 0.1978883296251297

step: 110, loss: 0.13496632874011993

step: 120, loss: 0.09040269255638123

step: 130, loss: 0.14535540342330933

step: 140, loss: 0.11177774518728256

```

step: 150, loss: 0.06419442594051361
step: 160, loss: 0.2359686642885208
step: 170, loss: 0.07893222570419312
step: 180, loss: 0.19089825451374054
step: 190, loss: 0.13609574735164642
step: 200, loss: 0.060646507889032364
step: 210, loss: 0.04124847427010536
step: 220, loss: 0.08228443562984467
step: 230, loss: 0.19135484099388123
step: 240, loss: 0.025132490321993828
step: 250, loss: 0.05047457665205002
step: 260, loss: 0.14507777988910675
step: 270, loss: 0.26718467473983765
step: 280, loss: 0.03094392642378807
step: 290, loss: 0.10588528215885162
step: 300, loss: 0.047462016344070435
step: 310, loss: 0.10413139313459396
step: 320, loss: 0.09352907538414001
step: 330, loss: 0.16173361241817474
step: 340, loss: 0.3397269546985626
step: 350, loss: 0.07089350372552872
step: 360, loss: 0.10348284244537354
step: 370, loss: 0.016571877524256706
step: 380, loss: 0.0950603187084198
step: 390, loss: 0.07929378002882004
step: 400, loss: 0.08896438032388687
step: 410, loss: 0.22672109305858612
step: 420, loss: 0.022235389798879623
step: 430, loss: 0.015347864478826523
step: 440, loss: 0.026100821793079376
step: 450, loss: 0.03744407370686531
step: 460, loss: 0.027018265798687935
step: 470, loss: 0.04501655325293541
step: 480, loss: 0.20566557347774506
step: 490, loss: 0.09507662057876587
step: 500, loss: 0.0050076767802238464
step: 510, loss: 0.02313411422073841
step: 520, loss: 0.09531094878911972
step: 530, loss: 0.08652176707983017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9336384439359268, f1=0.9412300683371299,
best_f1=0.9412300683371299

```

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

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

```
step: 0, loss: 0.060537561774253845  
step: 10, loss: 0.06584525853395462  
step: 20, loss: 0.025136582553386688  
step: 30, loss: 0.07496745139360428  
step: 40, loss: 0.014310292899608612  
step: 50, loss: 0.03252239152789116  
step: 60, loss: 0.012481719255447388  
step: 70, loss: 0.02657110057771206  
step: 80, loss: 0.008078977465629578  
step: 90, loss: 0.004413698334246874  
step: 100, loss: 0.11916148662567139  
step: 110, loss: 0.018759537488222122  
step: 120, loss: 0.04978003352880478  
step: 130, loss: 0.012982816435396671  
step: 140, loss: 0.0789470300078392  
step: 150, loss: 0.05489038676023483  
step: 160, loss: 0.04093712940812111  
step: 170, loss: 0.048560746014118195  
step: 180, loss: 0.006203044205904007  
step: 190, loss: 0.004898150451481342  
step: 200, loss: 0.12941093742847443  
step: 210, loss: 0.03267578035593033  
step: 220, loss: 0.0011165211908519268  
step: 230, loss: 0.05973311886191368  
step: 240, loss: 0.08771800249814987  
step: 250, loss: 0.02308378368616104  
step: 260, loss: 0.018504289910197258  
step: 270, loss: 0.010018670000135899  
step: 280, loss: 0.06117593124508858  
step: 290, loss: 0.0544043630361557  
step: 300, loss: 0.01461788360029459  
step: 310, loss: 0.08241619914770126  
step: 320, loss: 0.030960090458393097  
step: 330, loss: 0.05272983759641647  
step: 340, loss: 0.10565369576215744  
step: 350, loss: 0.002386999549344182  
step: 360, loss: 0.09209217876195908  
step: 370, loss: 0.009546251967549324  
step: 380, loss: 0.20179156959056854  
step: 390, loss: 0.004261044319719076  
step: 400, loss: 0.05559427663683891  
step: 410, loss: 0.043460503220558167  
step: 420, loss: 0.03424437344074249  
step: 430, loss: 0.45133280754089355  
step: 440, loss: 0.20068156719207764
```

```

step: 450, loss: 0.07854390889406204
step: 460, loss: 0.03478160500526428
step: 470, loss: 0.00962738785892725
step: 480, loss: 0.00265188398770988
step: 490, loss: 0.18795284628868103
step: 500, loss: 0.016165761277079582
step: 510, loss: 0.024437081068754196
step: 520, loss: 0.33113327622413635
step: 530, loss: 0.08120544254779816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9430004559963521, f1=0.9394629039599454,
best_f1=0.9394629039599454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.084659144282341
step: 10, loss: 0.04175718128681183
step: 20, loss: 0.005322752986103296
step: 30, loss: 0.041733771562576294
step: 40, loss: 0.04537458345293999
step: 50, loss: 0.01824982464313507
step: 60, loss: 0.0193234384059906
step: 70, loss: 0.025816785171628
step: 80, loss: 0.01426960714161396
step: 90, loss: 0.009815261699259281
step: 100, loss: 0.015225442126393318
step: 110, loss: 0.04970743507146835
step: 120, loss: 0.07252413779497147
step: 130, loss: 0.071133092045784
step: 140, loss: 0.021372506394982338
step: 150, loss: 0.014240846037864685
step: 160, loss: 0.012429825030267239
step: 170, loss: 0.04635930061340332
step: 180, loss: 0.0035892846062779427
step: 190, loss: 0.0008290195255540311
step: 200, loss: 0.02579624392092228
step: 210, loss: 0.03304202854633331
step: 220, loss: 0.14012816548347473
step: 230, loss: 0.0026422059163451195
step: 240, loss: 0.07312271744012833
step: 250, loss: 0.036287806928157806
step: 260, loss: 0.10244306921958923

```

```

step: 270, loss: 0.003147395793348551
step: 280, loss: 0.035907257348299026
step: 290, loss: 0.008836542256176472
step: 300, loss: 0.08702603727579117
step: 310, loss: 0.07672324031591415
step: 320, loss: 0.00793498381972313
step: 330, loss: 0.0015743611147627234
step: 340, loss: 0.004854653030633926
step: 350, loss: 0.04835209995508194
step: 360, loss: 0.004210443701595068
step: 370, loss: 0.03406129032373428
step: 380, loss: 0.005260044708848
step: 390, loss: 0.01160106249153614
step: 400, loss: 0.021815625950694084
step: 410, loss: 0.15378524363040924
step: 420, loss: 0.008596825413405895
step: 430, loss: 0.053246885538101196
step: 440, loss: 0.08381815999746323
step: 450, loss: 0.04215167090296745
step: 460, loss: 0.146370530128479
step: 470, loss: 0.028895366936922073
step: 480, loss: 0.05940739065408707
step: 490, loss: 0.03639484569430351
step: 500, loss: 0.022789079695940018
step: 510, loss: 0.028854291886091232
step: 520, loss: 0.003696009051054716
step: 530, loss: 0.028039084747433662
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.945287356321839, f1=0.9386446886446886,
best_f1=0.9386446886446886
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025300407782197
step: 10, loss: 0.004367470741271973
step: 20, loss: 0.008600112050771713
step: 30, loss: 0.0776984915137291
step: 40, loss: 0.0042799850925803185
step: 50, loss: 0.03213101625442505
step: 60, loss: 0.005186030641198158
step: 70, loss: 0.00535530736669898
step: 80, loss: 0.0052667236886918545

```

```
step: 90, loss: 0.08253943175077438
step: 100, loss: 0.0011242316104471684
step: 110, loss: 0.03772525116801262
step: 120, loss: 0.0009706601267680526
step: 130, loss: 0.04638735204935074
step: 140, loss: 0.07352264970541
step: 150, loss: 0.03565896674990654
step: 160, loss: 0.01736319251358509
step: 170, loss: 0.032152704894542694
step: 180, loss: 0.028611568734049797
step: 190, loss: 0.028175894170999527
step: 200, loss: 0.14567062258720398
step: 210, loss: 0.0008497737580910325
step: 220, loss: 0.019630996510386467
step: 230, loss: 0.11140759289264679
step: 240, loss: 0.04078876972198486
step: 250, loss: 0.08840128779411316
step: 260, loss: 0.0012541976757347584
step: 270, loss: 0.09699173271656036
step: 280, loss: 0.018158549442887306
step: 290, loss: 0.023414622992277145
step: 300, loss: 0.002711244858801365
step: 310, loss: 0.004783038515597582
step: 320, loss: 0.01969609037041664
step: 330, loss: 0.02754962258040905
step: 340, loss: 0.028209449723362923
step: 350, loss: 0.07295287400484085
step: 360, loss: 0.08255970478057861
step: 370, loss: 0.017032740637660027
step: 380, loss: 0.07053682953119278
step: 390, loss: 0.0008780521457083523
step: 400, loss: 0.01158726867288351
step: 410, loss: 0.002656960394233465
step: 420, loss: 0.0366181954741478
step: 430, loss: 0.017309391871094704
step: 440, loss: 0.1175636276602745
step: 450, loss: 0.03044453263282776
step: 460, loss: 0.01656354032456875
step: 470, loss: 0.005313632544130087
step: 480, loss: 0.003783345455303788
step: 490, loss: 0.00621027359738946
step: 500, loss: 0.04969693347811699
step: 510, loss: 0.026044340804219246
step: 520, loss: 0.010671399533748627
step: 530, loss: 0.15528735518455505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```



```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9553488372093023, f1=0.9502814258911819,
best_f1=0.9502814258911819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017405877588316798
step: 10, loss: 0.010901754721999168
step: 20, loss: 0.044954635202884674
step: 30, loss: 0.0029001024086028337
step: 40, loss: 0.02573912963271141
step: 50, loss: 0.16404476761817932
step: 60, loss: 0.0144600048661232
step: 70, loss: 0.0012325989082455635
step: 80, loss: 0.007266985718160868
step: 90, loss: 0.1362525373697281
step: 100, loss: 0.07637004554271698
step: 110, loss: 0.0023547776509076357
step: 120, loss: 0.1604951173067093
step: 130, loss: 0.011018206365406513
step: 140, loss: 0.006299784407019615
step: 150, loss: 0.025653202086687088
step: 160, loss: 0.0025220983661711216
step: 170, loss: 0.035396378487348557
step: 180, loss: 0.01295658852905035
step: 190, loss: 0.003738704603165388
step: 200, loss: 0.002308484399691224
step: 210, loss: 0.0030852914787828922
step: 220, loss: 0.039153024554252625
step: 230, loss: 0.0011371019063517451
step: 240, loss: 0.024115411564707756
step: 250, loss: 0.18570195138454437
step: 260, loss: 0.00047543051186949015
step: 270, loss: 0.0029778836760669947
step: 280, loss: 0.004144138190895319
step: 290, loss: 0.00201601255685091
step: 300, loss: 0.13383881747722626
step: 310, loss: 0.044776979833841324
step: 320, loss: 0.009517228230834007
step: 330, loss: 0.0012114958371967077
step: 340, loss: 0.0025974377058446407
step: 350, loss: 0.0029897866770625114
step: 360, loss: 0.00043016538256779313
step: 370, loss: 0.019180623814463615
step: 380, loss: 0.0005191625677980483

```

```

step: 390, loss: 0.012809574604034424
step: 400, loss: 0.005224583670496941
step: 410, loss: 0.06362749636173248
step: 420, loss: 0.17626264691352844
step: 430, loss: 0.020806942135095596
step: 440, loss: 0.011462636291980743
step: 450, loss: 0.015522628091275692
step: 460, loss: 0.038812071084976196
step: 470, loss: 0.032025452703237534
step: 480, loss: 0.005410696379840374
step: 490, loss: 0.00804110337048769
step: 500, loss: 0.017097916454076767
step: 510, loss: 0.000584652298130095
step: 520, loss: 0.03991403430700302
step: 530, loss: 0.05088851600885391
/usr/local/lib/python3.7/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=0.9510945505356311, f1=0.9418386491557222,
best_f1=0.9502814258911819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012670041993260384
step: 10, loss: 0.004168736282736063
step: 20, loss: 0.06293928623199463
step: 30, loss: 0.002238105284050107
step: 40, loss: 0.0004837955057155341
step: 50, loss: 0.001509870053268969
step: 60, loss: 0.0077648828737437725
step: 70, loss: 0.05500645935535431
step: 80, loss: 0.00013572018360719085
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.0038186036981642246
step: 100, loss: 0.0024120588786900043
step: 110, loss: 0.001994194695726037
step: 120, loss: 0.019624531269073486
step: 130, loss: 0.0024396178778260946
step: 140, loss: 0.002495997818186879
step: 150, loss: 0.0008329258416779339
step: 160, loss: 0.06889612227678299
step: 170, loss: 0.007399348542094231
step: 180, loss: 0.004618315491825342
step: 190, loss: 0.0031085798982530832
step: 200, loss: 0.018981782719492912
step: 210, loss: 0.00027683735243044794
step: 220, loss: 0.018302422016859055
step: 230, loss: 0.005280486773699522

```

```

step: 240, loss: 0.0005677388980984688
step: 250, loss: 0.02158469147980213
step: 260, loss: 0.0017960878321900964
step: 270, loss: 0.0003069759113714099
step: 280, loss: 0.005368090234696865
step: 290, loss: 0.0208982415497303
step: 300, loss: 0.009745133109390736
step: 310, loss: 0.01462134625762701
step: 320, loss: 8.310218254337087e-05
step: 330, loss: 0.004950202535837889
step: 340, loss: 0.0007071794243529439
step: 350, loss: 0.01154258567839861
step: 360, loss: 0.15884849429130554
step: 370, loss: 0.008377904072403908
step: 380, loss: 0.002935313619673252
step: 390, loss: 0.014397923834621906
step: 400, loss: 0.006094672251492739
step: 410, loss: 0.00022916072339285165
step: 420, loss: 0.0009734079940244555
step: 430, loss: 0.0006836270913481712
step: 440, loss: 0.007554224226623774
step: 450, loss: 0.1955416202545166
step: 460, loss: 0.001122316694818437
step: 470, loss: 0.009786962531507015
step: 480, loss: 0.006118998862802982
step: 490, loss: 0.011066392995417118
step: 500, loss: 0.010569209232926369
step: 510, loss: 0.1740085631608963
step: 520, loss: 0.0005898531526327133
step: 530, loss: 0.0018575657159090042
/usr/local/lib/python3.7/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=0.9509848831882729, f1=0.9418181818181819,
best_f1=0.9502814258911819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018455061363056302
step: 10, loss: 0.010222391225397587
step: 20, loss: 0.0024404472205787897
step: 30, loss: 0.07097327709197998
step: 40, loss: 0.013028084300458431
step: 50, loss: 0.011630665510892868
step: 60, loss: 0.001285871141590178
step: 70, loss: 0.00047916348557919264
step: 80, loss: 0.0004320019797887653
step: 90, loss: 0.0020758507307618856

```

```

step: 100, loss: 0.00516851618885994
step: 110, loss: 0.0007679623086005449
step: 120, loss: 0.0005288478569127619
step: 130, loss: 0.0001176316145574674
step: 140, loss: 0.00018829788314178586
step: 150, loss: 0.0001780871389200911
step: 160, loss: 7.561050733784214e-05
step: 170, loss: 0.003937148954719305
step: 180, loss: 0.03767912834882736
step: 190, loss: 0.02120625227689743
step: 200, loss: 0.0003896873677149415
step: 210, loss: 0.00021427989122457802
step: 220, loss: 0.017784731462597847
step: 230, loss: 0.00016094028251245618
step: 240, loss: 0.0004617770609911531
step: 250, loss: 0.0002524334122426808
step: 260, loss: 0.007808547466993332
step: 270, loss: 0.0008036249200813472
step: 280, loss: 0.012439284473657608
step: 290, loss: 0.0009824794251471758
step: 300, loss: 0.003964161034673452
step: 310, loss: 0.0007431551930494606
step: 320, loss: 0.004282419104129076
step: 330, loss: 0.004848862066864967
step: 340, loss: 0.006165707018226385
step: 350, loss: 0.0021531691309064627
step: 360, loss: 0.009146949276328087
step: 370, loss: 0.00937998853623867
step: 380, loss: 0.012248107232153416
step: 390, loss: 0.0009571159025654197
step: 400, loss: 0.023706186562776566
step: 410, loss: 0.0015581881161779165
step: 420, loss: 0.18298570811748505
step: 430, loss: 0.010507875122129917
step: 440, loss: 0.042452648282051086
step: 450, loss: 0.005487675778567791
step: 460, loss: 0.0008793942397460341
step: 470, loss: 0.1453639715909958
step: 480, loss: 0.006606228183954954
step: 490, loss: 0.0026016789488494396
step: 500, loss: 0.0007242754218168557
step: 510, loss: 0.0006136340089142323
step: 520, loss: 0.0010505208047106862
step: 530, loss: 0.0013217503437772393
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9578508568781844, f1=0.9506976744186046,

best_f1=0.9506976744186046

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

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

step: 0, loss: 0.12774385511875153

step: 10, loss: 0.0017984297592192888

step: 20, loss: 0.0005490533076226711

step: 30, loss: 0.0006174108712002635

step: 40, loss: 0.00016974138270597905

step: 50, loss: 0.004477457143366337

step: 60, loss: 0.0005629587103612721

step: 70, loss: 0.09170432388782501

step: 80, loss: 0.0006511911633424461

step: 90, loss: 0.0002080878766719252

step: 100, loss: 0.0005764125380665064

step: 110, loss: 0.001927815144881606

step: 120, loss: 0.0005614808760583401

step: 130, loss: 0.0016618514200672507

step: 140, loss: 0.0003992717538494617

step: 150, loss: 0.0003014236281160265

step: 160, loss: 0.0005197149002924562

step: 170, loss: 0.14085182547569275

step: 180, loss: 0.0016870938707143068

step: 190, loss: 0.0032029012218117714

step: 200, loss: 0.001507273525930941

step: 210, loss: 0.11056423932313919

step: 220, loss: 0.0003606526006478816

step: 230, loss: 0.04907093569636345

step: 240, loss: 0.009314426220953465

step: 250, loss: 0.0025711350608617067

step: 260, loss: 0.0005829564179293811

step: 270, loss: 0.03983057662844658

step: 280, loss: 0.0018185467924922705

step: 290, loss: 0.00042385130655020475

step: 300, loss: 0.00048238661838695407

step: 310, loss: 0.016947757452726364

step: 320, loss: 0.0008848885772749782

step: 330, loss: 0.0003510899841785431

step: 340, loss: 0.006977032404392958

step: 350, loss: 0.0002524942683521658

step: 360, loss: 0.024933097884058952

step: 370, loss: 0.05308195948600769

step: 380, loss: 0.00032205029856413603

step: 390, loss: 0.005573611706495285

```

step: 400, loss: 0.0017306520603597164
step: 410, loss: 0.0003378749534022063
step: 420, loss: 0.0010191269684582949
step: 430, loss: 0.0007925716927275062
step: 440, loss: 0.0008555416134186089
step: 450, loss: 0.0006015099352225661
step: 460, loss: 0.00046804134035483
step: 470, loss: 0.02786104567348957
step: 480, loss: 0.00024092000967357308
step: 490, loss: 0.0010390611132606864
step: 500, loss: 0.0004238744731992483
step: 510, loss: 0.0013549806317314506
step: 520, loss: 0.00035590873449109495
step: 530, loss: 0.012988326139748096
/usr/local/lib/python3.7/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=0.9556470050297211, f1=0.9510745313214449,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.520814789226279e-05
step: 10, loss: 0.0006015971885062754
step: 20, loss: 0.00034468539524823427
step: 30, loss: 0.11883768439292908
step: 40, loss: 0.0009768652962520719
step: 50, loss: 0.012035920284688473
step: 60, loss: 0.0009025884792208672
step: 70, loss: 0.04268893972039223
step: 80, loss: 0.05864904075860977
step: 90, loss: 0.01912032440304756
step: 100, loss: 0.0009643882513046265
step: 110, loss: 0.03966251015663147
step: 120, loss: 0.0003750855103135109
step: 130, loss: 0.00127210165373981
step: 140, loss: 0.00031996279722079635
step: 150, loss: 0.07295391708612442
step: 160, loss: 0.0006502957548946142
step: 170, loss: 0.0036596297286450863
step: 180, loss: 0.0007599233649671078
step: 190, loss: 0.00017663135076873004
step: 200, loss: 0.001090813777409494
step: 210, loss: 0.0016376241110265255
step: 220, loss: 0.0003473003744147718
step: 230, loss: 0.0007210436160676181
step: 240, loss: 0.0010961714433506131
step: 250, loss: 0.0006266772979870439

```

```

step: 260, loss: 0.0005019084201194346
step: 270, loss: 0.007546433247625828
step: 280, loss: 0.00032309486414305866
step: 290, loss: 0.00018442823784425855
step: 300, loss: 0.003755284706130624
step: 310, loss: 0.10480712354183197
step: 320, loss: 0.0001977054780581966
step: 330, loss: 0.0008041657274588943
step: 340, loss: 0.0012265853583812714
step: 350, loss: 0.006706492509692907
step: 360, loss: 0.00020356214372441173
step: 370, loss: 0.0009275751654058695
step: 380, loss: 0.006805552169680595
step: 390, loss: 6.793843931518495e-05
step: 400, loss: 0.0019887175876647234
step: 410, loss: 0.00031018638401292264
step: 420, loss: 0.00015535425336565822
step: 430, loss: 0.005875525530427694
step: 440, loss: 0.00021344881679397076
step: 450, loss: 0.01088421605527401
step: 460, loss: 8.631362288724631e-05
step: 470, loss: 0.0012419214472174644
step: 480, loss: 4.6672270400449634e-05
step: 490, loss: 0.00022576515038963407
step: 500, loss: 0.00023450347362086177
step: 510, loss: 0.003294369438663125
step: 520, loss: 0.00408233841881156
step: 530, loss: 0.008123720996081829
/usr/local/lib/python3.7/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=0.9502304147465437, f1=0.9463955637707948,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003174307756125927
step: 10, loss: 0.0010540231596678495
step: 20, loss: 7.549786096205935e-05
step: 30, loss: 0.0008939084364101291
step: 40, loss: 0.003435238031670451
step: 50, loss: 0.001176894991658628
step: 60, loss: 0.006228719837963581
step: 70, loss: 0.0001054018794093281
step: 80, loss: 0.0025415890850126743
step: 90, loss: 0.0008253917330875993
step: 100, loss: 0.0007441712077707052
step: 110, loss: 0.04201671853661537

```

```

step: 120, loss: 0.005181348416954279
step: 130, loss: 0.0017374573508277535
step: 140, loss: 0.00014526369341183454
step: 150, loss: 0.000414066860685125
step: 160, loss: 0.02567536197602749
step: 170, loss: 0.00020398605556692928
step: 180, loss: 0.013590787537395954
step: 190, loss: 3.11434268951416e-05
step: 200, loss: 4.081551014678553e-05
step: 210, loss: 0.12190936505794525
step: 220, loss: 0.0012608858523890376
step: 230, loss: 0.0029061681125313044
step: 240, loss: 0.0024314643815159798
step: 250, loss: 0.004909933544695377
step: 260, loss: 0.009694904088973999
step: 270, loss: 0.010831745341420174
step: 280, loss: 0.009673647582530975
step: 290, loss: 0.0014652636600658298
step: 300, loss: 0.017392370849847794
step: 310, loss: 0.02192460559308529
step: 320, loss: 0.025689462199807167
step: 330, loss: 0.0018867958569899201
step: 340, loss: 0.0008002996328286827
step: 350, loss: 0.000265954906353727
step: 360, loss: 0.00010732026566984132
step: 370, loss: 0.0005419024964794517
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 380, loss: 0.0008655795827507973
step: 390, loss: 0.0011616076808422804
step: 400, loss: 0.0006875368417240679
step: 410, loss: 0.0001477658370276913
step: 420, loss: 5.745899034081958e-05
step: 430, loss: 0.00012080423766747117
step: 440, loss: 0.00021814070350956172
step: 450, loss: 0.004784659016877413
step: 460, loss: 0.00039118423592299223
step: 470, loss: 0.0013839374296367168
step: 480, loss: 0.0001871580898296088
step: 490, loss: 0.1406112164258957
step: 500, loss: 0.12831620872020721
step: 510, loss: 6.422459409805015e-05
step: 520, loss: 0.011737310327589512
step: 530, loss: 0.01286434754729271
/usr/local/lib/python3.7/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=0.9556696220251982, f1=0.947022972339428,
best_f1=0.9506976744186046

```


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

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

step: 0, loss: 0.00035775857395492494
step: 10, loss: 0.00015616636665072292
step: 20, loss: 0.009540782310068607
step: 30, loss: 0.013069787062704563
step: 40, loss: 8.689276000950485e-05
step: 50, loss: 9.252134623238817e-05
step: 60, loss: 0.00023996626259759068
step: 70, loss: 0.003917198162525892
step: 80, loss: 0.0016918187029659748
step: 90, loss: 0.000456829322502017
step: 100, loss: 0.00011483347043395042
step: 110, loss: 8.615668048150837e-05
step: 120, loss: 0.0023033409379422665
step: 130, loss: 8.296513988170773e-05
step: 140, loss: 4.532166349235922e-05
step: 150, loss: 0.0001114486949518323
step: 160, loss: 0.0001283140736632049
step: 170, loss: 0.0004244203446432948
step: 180, loss: 2.504432814021129e-05
step: 190, loss: 0.0005442832480184734
step: 200, loss: 2.6779713152791373e-05
step: 210, loss: 3.918715810868889e-05
step: 220, loss: 0.006706495303660631
step: 230, loss: 0.00011111926141893491
step: 240, loss: 0.006215078756213188
step: 250, loss: 0.000818825326859951
step: 260, loss: 0.005461352877318859
step: 270, loss: 0.0012192273279652
step: 280, loss: 0.002263531554490328
step: 290, loss: 0.002207803539931774
step: 300, loss: 9.275455522583798e-05
step: 310, loss: 0.006844519637525082
step: 320, loss: 0.00013726871111430228
step: 330, loss: 1.738180981192272e-05
step: 340, loss: 0.0033705804962664843
step: 350, loss: 0.004262572154402733
step: 360, loss: 8.798505587037653e-05
step: 370, loss: 0.002705511637032032
step: 380, loss: 0.0012730566086247563
step: 390, loss: 0.0006463129539042711
step: 400, loss: 8.87379064806737e-05
step: 410, loss: 0.002792470157146454
step: 420, loss: 0.001109277131035924
step: 430, loss: 0.0012649609707295895
step: 440, loss: 0.00010029740224126726

```

step: 450, loss: 0.0008099405094981194
step: 460, loss: 0.0024228389374911785
step: 470, loss: 7.440961053362116e-05
step: 480, loss: 9.566758672008291e-05
step: 490, loss: 2.2946927856537513e-05
step: 500, loss: 0.001591412932612002
step: 510, loss: 0.0008625576738268137
step: 520, loss: 0.0016648174496367574
step: 530, loss: 0.008282499387860298
/usr/local/lib/python3.7/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=0.9536178107606679, f1=0.951048951048951,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00044056697515770793
step: 10, loss: 0.0001230748457601294
step: 20, loss: 0.001426956383511424
step: 30, loss: 5.689150566468015e-05
step: 40, loss: 0.0006528238300234079
step: 50, loss: 5.1168157369829714e-05
step: 60, loss: 9.29223679122515e-05
step: 70, loss: 0.0004679207340814173
step: 80, loss: 0.0029810776468366385
step: 90, loss: 0.0004830558318644762
step: 100, loss: 0.0462331548333168
step: 110, loss: 1.942685958056245e-05
step: 120, loss: 0.00042593051330186427
step: 130, loss: 7.064205419737846e-05
step: 140, loss: 0.00022496122983284295
step: 150, loss: 0.00014847585407551378
step: 160, loss: 0.00026846188120543957
step: 170, loss: 5.3360097808763385e-05
step: 180, loss: 1.644679286982864e-05
step: 190, loss: 0.0001279570278711617
step: 200, loss: 0.0003376344684511423
step: 210, loss: 6.274917541304603e-05
step: 220, loss: 7.571507740067318e-05
step: 230, loss: 0.0001289605861529708
step: 240, loss: 0.00021098324214108288
step: 250, loss: 9.475352999288589e-05
step: 260, loss: 4.869155236519873e-05
step: 270, loss: 5.584876998909749e-05
step: 280, loss: 4.356893987278454e-05
step: 290, loss: 0.00020678139117080718
step: 300, loss: 0.0004070914292242378

```

```

step: 310, loss: 0.0006261206581257284
step: 320, loss: 0.001902377582155168
step: 330, loss: 0.00042156901326961815
step: 340, loss: 9.22039762372151e-05
step: 350, loss: 2.1403331629699096e-05
step: 360, loss: 0.00014741311315447092
step: 370, loss: 2.8015732823405415e-05
step: 380, loss: 0.0002038756647380069
step: 390, loss: 0.004082426894456148
step: 400, loss: 3.091127655352466e-05
step: 410, loss: 0.000389215158065781
step: 420, loss: 0.001141752116382122
step: 430, loss: 0.0011044531129300594
step: 440, loss: 0.00019764964235946536
step: 450, loss: 0.0007262644357979298
step: 460, loss: 4.348847141955048e-05
step: 470, loss: 0.010550957173109055
step: 480, loss: 4.85358941659797e-05
step: 490, loss: 2.3624988898518495e-05
step: 500, loss: 2.5604505935916677e-05
step: 510, loss: 0.028689945116639137
step: 520, loss: 0.00022040570911485702
step: 530, loss: 4.36693626397755e-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=0.9544626593806922, f1=0.9422989550204453,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005776646081358194
step: 10, loss: 0.0004981953534297645
step: 20, loss: 0.0008886252762749791
step: 30, loss: 2.617024802020751e-05
step: 40, loss: 0.0012692911550402641
step: 50, loss: 0.000677438045386225
step: 60, loss: 0.0004966820706613362
step: 70, loss: 0.001156217185780406
step: 80, loss: 0.00019302540749777108
step: 90, loss: 5.0060494686476886e-05
step: 100, loss: 1.6323559975717217e-05
step: 110, loss: 0.0007288388442248106
step: 120, loss: 9.210097778122872e-05
step: 130, loss: 0.0001498455967521295
step: 140, loss: 8.027088915696368e-05
step: 150, loss: 0.00034163586678914726
step: 160, loss: 0.00032262911554425955

```

```

step: 170, loss: 0.0014826747355982661
step: 180, loss: 3.4750632039504126e-05
step: 190, loss: 0.002303770510479808
step: 200, loss: 4.283002272131853e-05
step: 210, loss: 2.51164074143162e-05
step: 220, loss: 0.000505105301272124
step: 230, loss: 0.0072924294508993626
step: 240, loss: 0.00023416793555952609
step: 250, loss: 1.7024278349708766e-05
step: 260, loss: 3.0705381504958495e-05
step: 270, loss: 0.00015822469140402973
step: 280, loss: 4.8852863983483985e-05
step: 290, loss: 1.2155426702520344e-05
step: 300, loss: 1.8514165276428685e-05
step: 310, loss: 0.00014441994426306337
step: 320, loss: 0.0003836637479253113
step: 330, loss: 1.7396807379554957e-05
step: 340, loss: 0.0003749714232981205
step: 350, loss: 9.488197065365966e-06
step: 360, loss: 0.1042945459485054
step: 370, loss: 0.01046940591186285
step: 380, loss: 1.1350744898663834e-05
step: 390, loss: 0.00010596077481750399
step: 400, loss: 0.001105748233385384
step: 410, loss: 2.4570217647124082e-05
step: 420, loss: 0.00949022825807333
step: 430, loss: 0.0003230785659980029
step: 440, loss: 1.372726183035411e-05
step: 450, loss: 1.785855056368746e-05
step: 460, loss: 0.0005765100941061974
step: 470, loss: 0.0003132903075311333
step: 480, loss: 7.4989516178902704e-06
step: 490, loss: 2.0737468730658293e-05
step: 500, loss: 9.196022438118234e-05
step: 510, loss: 8.696984150446951e-05
step: 520, loss: 0.00011206357885384932
step: 530, loss: 0.002086672466248274
/usr/local/lib/python3.7/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=0.9541454377026402, f1=0.9506057781919852,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.316562262829393e-05
step: 10, loss: 0.00012477763812057674
step: 20, loss: 9.746552677825093e-05

```

step: 30, loss: 0.006108209956437349
step: 40, loss: 3.343632124597207e-05
step: 50, loss: 0.00032309413654729724
step: 60, loss: 2.2060117771616206e-05
step: 70, loss: 8.444109698757529e-05
step: 80, loss: 6.459971336880699e-05
step: 90, loss: 1.0222039236396085e-05
step: 100, loss: 9.132848936133087e-05
step: 110, loss: 0.0003457230923231691
step: 120, loss: 6.823646981501952e-05
step: 130, loss: 2.3940805476740934e-05
step: 140, loss: 9.931702516041696e-05
step: 150, loss: 5.034155401517637e-05
step: 160, loss: 1.8484284737496637e-05
step: 170, loss: 0.007299392484128475
step: 180, loss: 2.6700467060436495e-05
step: 190, loss: 4.180723408353515e-05
step: 200, loss: 1.5869332855800167e-05
step: 210, loss: 4.5925291487947106e-05
step: 220, loss: 1.0866508091567084e-05
step: 230, loss: 1.1633940630417783e-05
step: 240, loss: 0.00017642887542024255
step: 250, loss: 3.912904503522441e-05
step: 260, loss: 2.563833368185442e-05
step: 270, loss: 9.098667214857414e-05
step: 280, loss: 8.63333189045079e-05
step: 290, loss: 1.6256410162895918e-05
step: 300, loss: 5.528199108084664e-05
step: 310, loss: 0.0011391851585358381
step: 320, loss: 4.379640085971914e-05
step: 330, loss: 0.0004749361833091825
step: 340, loss: 0.00023206007608678192
step: 350, loss: 9.96876769931987e-06
step: 360, loss: 6.732048495905474e-05
step: 370, loss: 2.5522977011860348e-05
step: 380, loss: 0.00033694683224894106
step: 390, loss: 0.00014182056474965066
step: 400, loss: 0.012335538864135742
step: 410, loss: 0.0024413932114839554
step: 420, loss: 4.5206950744614005e-05
step: 430, loss: 2.66928782366449e-05
step: 440, loss: 2.855607272067573e-05
step: 450, loss: 5.673485793522559e-05
step: 460, loss: 0.009596713818609715
step: 470, loss: 1.5221215107885655e-05
step: 480, loss: 5.720699846278876e-05
step: 490, loss: 2.6901601813733578e-05
step: 500, loss: 3.232738527003676e-05

```

step: 510, loss: 0.0007203956483863294
step: 520, loss: 3.911030580638908e-05
step: 530, loss: 0.0005068924510851502
/usr/local/lib/python3.7/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=0.9512195121951219, f1=0.9480093676814988,
best_f1=0.9506976744186046
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.728074494167231e-05
step: 10, loss: 0.0001709338539512828
step: 20, loss: 3.707108771777712e-05
step: 30, loss: 0.0035558410454541445
step: 40, loss: 1.9854967831633985e-05
step: 50, loss: 0.00018192357674706727
step: 60, loss: 0.00040360333514399827
step: 70, loss: 4.013353100162931e-05
step: 80, loss: 2.010052776313387e-05
step: 90, loss: 4.224647273076698e-05
step: 100, loss: 8.22163656266639e-06
step: 110, loss: 2.4987743017845787e-05
step: 120, loss: 1.2092046745237894e-05
step: 130, loss: 1.0561057933955453e-05
step: 140, loss: 0.0013207648880779743
step: 150, loss: 1.8580720279715024e-05
step: 160, loss: 1.1749344594136346e-05
step: 170, loss: 7.02211173120304e-06
step: 180, loss: 1.6703783330740407e-05
step: 190, loss: 0.0009456276311539114
step: 200, loss: 0.0002538894477766007
step: 210, loss: 7.651688065379858e-06
step: 220, loss: 1.9407892978051677e-05
step: 230, loss: 0.003382929367944598
step: 240, loss: 1.3004667380300816e-05
step: 250, loss: 7.115252628864255e-06
step: 260, loss: 8.631404853076674e-06
step: 270, loss: 1.5217487089103088e-05
step: 280, loss: 2.5602812456781976e-05
step: 290, loss: 4.480919233174063e-05
step: 300, loss: 1.0322616617486347e-05
step: 310, loss: 0.016657063737511635
step: 320, loss: 2.0997833416913636e-05
step: 330, loss: 2.0033930923091248e-05
step: 340, loss: 1.4148377886158414e-05
step: 350, loss: 0.0014104762813076377
step: 360, loss: 1.0207197192357853e-05

```

```

step: 370, loss: 0.0009499578736722469
step: 380, loss: 1.2188939763291273e-05
step: 390, loss: 8.018040534807369e-05
step: 400, loss: 0.0011889482848346233
step: 410, loss: 1.7757622117642313e-05
step: 420, loss: 1.0121458217327017e-05
step: 430, loss: 6.686848337267293e-06
step: 440, loss: 0.00011126508616143838
step: 450, loss: 0.00061385816661641
step: 460, loss: 0.00027449597837403417
step: 470, loss: 1.1376851034583524e-05
step: 480, loss: 0.007639298215508461
step: 490, loss: 1.8088614524458535e-05
step: 500, loss: 0.0005426345160230994
step: 510, loss: 1.76454250322422e-05
step: 520, loss: 1.7235921404790133e-05
step: 530, loss: 7.5585426202451345e-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=0.9524705882352942, f1=0.9487058823529413,
best_f1=0.9506976744186046
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
5742it [00:34, 165.60it/s]
load_f1 = 0.9584487534626038
real_f1 = 0.9552789303826649
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, 141.31it/s]
```

3 DK

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

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 446 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: 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: 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: 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: 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: 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: 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: 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: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: 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: 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: 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: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=a5d27351d9d83ea421ed8c5122fa3822f53c8db05afa00a49647def52859e4df
  Stored in directory: /tmp/pip-ephem-wheel-cache-z32hkwc3/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 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: 0.43853822350502014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2828282828282828, f1=0.2772277227722772,
best_f1=0.2772277227722772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46107572317123413
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.3333333333333333, f1=0.32432432432432434,
best_f1=0.32432432432432434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4124232530593872
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 3: dev_f1=0.41935483870967744, f1=0.36065573770491804,
best_f1=0.36065573770491804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26131004095077515
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285714, f1=0.35555555555555557,
best_f1=0.35555555555555557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2924952208995819
/usr/local/lib/python3.7/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=0.40625, f1=0.4067796610169491, best_f1=0.35555555555555557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3666001260280609
/usr/local/lib/python3.7/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=0.56, f1=0.5306122448979592, best_f1=0.3555555555555557
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4646848440170288
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.5714285714285714,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2582986354827881
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8461538461538461, f1=0.8571428571428571,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23933638632297516
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.8387096774193549, f1=0.7058823529411764,
best_f1=0.8571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21434253454208374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8666666666666666, f1=0.7741935483870968,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10406529903411865
/usr/local/lib/python3.7/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=0.8571428571428571, f1=0.8888888888888889,
best_f1=0.7741935483870968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014185910113155842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8888888888888889, f1=0.8888888888888889,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005562989506870508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8387096774193549, f1=0.7647058823529412,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005445131100714207
/usr/local/lib/python3.7/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=0.7647058823529412,
best_f1=0.8888888888888889
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005315236747264862
/usr/local/lib/python3.7/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=0.7647058823529412,
best_f1=0.888888888888889
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 122530.23it/s]
/usr/local/lib/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 = 0.8275862068965518
real_f1 = 0.7333333333333334
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, 140.19it/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: 0.5646217465400696
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: 0.5094962120056152
step: 20, loss: 0.45888692140579224
step: 30, loss: 0.28208446502685547
step: 40, loss: 0.3386639952659607
step: 50, loss: 0.7403507232666016
step: 60, loss: 0.4777058959007263
step: 70, loss: 0.4356921911239624
step: 80, loss: 0.5974984765052795
step: 90, loss: 0.4372307360172272
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.4834124445915222
step: 110, loss: 0.5710242986679077
step: 120, loss: 0.4429129958152771
step: 130, loss: 0.32882410287857056
step: 140, loss: 0.399247407913208
step: 150, loss: 0.5684493780136108
step: 160, loss: 0.48572850227355957
step: 170, loss: 0.4768447279930115
step: 180, loss: 0.4031793475151062
step: 190, loss: 0.2751653492450714
step: 200, loss: 0.4093119204044342
step: 210, loss: 0.468589723110199
step: 220, loss: 0.6603602170944214
step: 230, loss: 0.4401867091655731
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5016226768493652

```

```

step: 10, loss: 0.4971943795681
step: 20, loss: 0.5971491932868958
step: 30, loss: 0.45910540223121643
step: 40, loss: 0.58409184217453
step: 50, loss: 0.49324408173561096
step: 60, loss: 0.47118061780929565
step: 70, loss: 0.5377886295318604
step: 80, loss: 0.5544477105140686
step: 90, loss: 0.4955502152442932
step: 100, loss: 0.5684638023376465
step: 110, loss: 0.43817251920700073
step: 120, loss: 0.5951164364814758
step: 130, loss: 0.43740391731262207
step: 140, loss: 0.5382401347160339
step: 150, loss: 0.5501158237457275
step: 160, loss: 0.5413830876350403
step: 170, loss: 0.24484772980213165
step: 180, loss: 0.7254502177238464
step: 190, loss: 0.5135303735733032
step: 200, loss: 0.6707589626312256
step: 210, loss: 0.4935600757598877
step: 220, loss: 0.35222840309143066
step: 230, loss: 0.39846673607826233
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4481523931026459
step: 10, loss: 0.542455792427063
step: 20, loss: 0.4531475901603699
step: 30, loss: 0.450580358505249
step: 40, loss: 0.5252756476402283
step: 50, loss: 0.47374746203422546
step: 60, loss: 0.6024888753890991
step: 70, loss: 0.42061272263526917
step: 80, loss: 0.34355753660202026
step: 90, loss: 0.45399725437164307
step: 100, loss: 0.49844983220100403
step: 110, loss: 0.522023618221283
step: 120, loss: 0.27825647592544556
step: 130, loss: 0.5997352004051208
step: 140, loss: 0.38411983847618103
step: 150, loss: 0.42197754979133606
step: 160, loss: 0.6125333309173584

```

```

step: 170, loss: 0.33561113476753235
step: 180, loss: 0.6848108768463135
step: 190, loss: 0.6674513816833496
step: 200, loss: 0.6096991300582886
step: 210, loss: 0.3364628851413727
step: 220, loss: 0.6889395713806152
step: 230, loss: 0.3696596920490265
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5186493396759033
step: 10, loss: 0.39238855242729187
step: 20, loss: 0.4438052475452423
step: 30, loss: 0.32705792784690857
step: 40, loss: 0.32532548904418945
step: 50, loss: 0.4811129868030548
step: 60, loss: 0.5180948972702026
step: 70, loss: 0.3900316655635834
step: 80, loss: 0.49020859599113464
step: 90, loss: 0.5549798607826233
step: 100, loss: 0.4219062328338623
step: 110, loss: 0.37014293670654297
step: 120, loss: 0.557937741279602
step: 130, loss: 0.37515029311180115
step: 140, loss: 0.48637714982032776
step: 150, loss: 0.3787688612937927
step: 160, loss: 0.5784245133399963
step: 170, loss: 0.5444355607032776
step: 180, loss: 0.38992443680763245
step: 190, loss: 0.5198519825935364
step: 200, loss: 0.6127853989601135
step: 210, loss: 0.49751177430152893
step: 220, loss: 0.37489816546440125
step: 230, loss: 0.44289350509643555
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32949599623680115

```

```

step: 10, loss: 0.7340983152389526
step: 20, loss: 0.5595385432243347
step: 30, loss: 0.5723679065704346
step: 40, loss: 0.5547119975090027
step: 50, loss: 0.2518369257450104
step: 60, loss: 0.3221193850040436
step: 70, loss: 0.3320111930370331
step: 80, loss: 0.4752531945705414
step: 90, loss: 0.6603960990905762
step: 100, loss: 0.350383996963501
step: 110, loss: 0.3471892178058624
step: 120, loss: 0.3945341408252716
step: 130, loss: 0.4230586886405945
step: 140, loss: 0.48454388976097107
step: 150, loss: 0.48625874519348145
step: 160, loss: 0.38761255145072937
step: 170, loss: 0.485821008682251
step: 180, loss: 0.442817360162735
step: 190, loss: 0.605762779712677
step: 200, loss: 0.6390520930290222
step: 210, loss: 0.6644774675369263
step: 220, loss: 0.5053000450134277
step: 230, loss: 0.5745708346366882
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3794116973876953
step: 10, loss: 0.6406263113021851
step: 20, loss: 0.613610565662384
step: 30, loss: 0.4471483826637268
step: 40, loss: 0.3921225965023041
step: 50, loss: 0.44631531834602356
step: 60, loss: 0.4805990159511566
step: 70, loss: 0.34473156929016113
step: 80, loss: 0.6182169914245605
step: 90, loss: 0.5408122539520264
step: 100, loss: 0.4889290928840637
step: 110, loss: 0.5347052812576294
step: 120, loss: 0.4969916343688965
step: 130, loss: 0.44672608375549316
step: 140, loss: 0.4331759512424469
step: 150, loss: 0.24531058967113495
step: 160, loss: 0.5306875109672546

```

```

step: 170, loss: 0.5291431546211243
step: 180, loss: 0.33513253927230835
step: 190, loss: 0.5380997657775879
step: 200, loss: 0.6805984973907471
step: 210, loss: 0.2895033657550812
step: 220, loss: 0.3753299415111542
step: 230, loss: 0.4836558401584625
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6100184321403503
step: 10, loss: 0.49237972497940063
step: 20, loss: 0.3380649983882904
step: 30, loss: 0.27450934052467346
step: 40, loss: 0.48368361592292786
step: 50, loss: 0.435806542634964
step: 60, loss: 0.4372462332248688
step: 70, loss: 0.4760500490665436
step: 80, loss: 0.527686357498169
step: 90, loss: 0.4300341308116913
step: 100, loss: 0.48492148518562317
step: 110, loss: 0.5435153245925903
step: 120, loss: 0.5312885046005249
step: 130, loss: 0.48651042580604553
step: 140, loss: 0.3340567946434021
step: 150, loss: 0.5889121294021606
step: 160, loss: 0.39406755566596985
step: 170, loss: 0.4520301818847656
step: 180, loss: 0.3921491503715515
step: 190, loss: 0.5260653495788574
step: 200, loss: 0.3925630450248718
step: 210, loss: 0.4275943636894226
step: 220, loss: 0.8827321529388428
step: 230, loss: 0.521414041519165
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4322030544281006

```

```

step: 10, loss: 0.47832703590393066
step: 20, loss: 0.4362843334674835
step: 30, loss: 0.43062853813171387
step: 40, loss: 0.5334702134132385
step: 50, loss: 0.43473780155181885
step: 60, loss: 0.513399600982666
step: 70, loss: 0.40771377086639404
step: 80, loss: 0.429184228181839
step: 90, loss: 0.4701439142227173
step: 100, loss: 0.512769877910614
step: 110, loss: 0.4408678412437439
step: 120, loss: 0.4353569447994232
step: 130, loss: 0.5411547422409058
step: 140, loss: 0.38173791766166687
step: 150, loss: 0.5357998013496399
step: 160, loss: 0.44417697191238403
step: 170, loss: 0.44293731451034546
step: 180, loss: 0.33297571539878845
step: 190, loss: 0.39482995867729187
step: 200, loss: 0.29126399755477905
step: 210, loss: 0.6619981527328491
step: 220, loss: 0.4784253239631653
step: 230, loss: 0.4361852705478668
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37809282541275024
step: 10, loss: 0.43455949425697327
step: 20, loss: 0.433003693819046
step: 30, loss: 0.4849836230278015
step: 40, loss: 0.44029438495635986
step: 50, loss: 0.48634499311447144
step: 60, loss: 0.3869723975658417
step: 70, loss: 0.39075541496276855
step: 80, loss: 0.30911681056022644
step: 90, loss: 0.49049270153045654
step: 100, loss: 0.48555120825767517
step: 110, loss: 0.43013209104537964
step: 120, loss: 0.3873218297958374
step: 130, loss: 0.572712779045105
step: 140, loss: 0.5233452916145325
step: 150, loss: 0.48417213559150696
step: 160, loss: 0.7041139602661133

```

```

step: 170, loss: 0.3757241666316986
step: 180, loss: 0.47636106610298157
step: 190, loss: 0.39433637261390686
step: 200, loss: 0.42842987179756165
step: 210, loss: 0.5927700996398926
step: 220, loss: 0.42868489027023315
step: 230, loss: 0.34603017568588257
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.434992253780365
step: 10, loss: 0.619939923286438
step: 20, loss: 0.48916542530059814
step: 30, loss: 0.3450515568256378
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.7817617654800415
step: 50, loss: 0.27536866068840027
step: 60, loss: 0.435967355966568
step: 70, loss: 0.48990216851234436
step: 80, loss: 0.32288825511932373
step: 90, loss: 0.3380865454673767
step: 100, loss: 0.3915032744407654
step: 110, loss: 0.3944149315357208
step: 120, loss: 0.38519275188446045
step: 130, loss: 0.4849603474140167
step: 140, loss: 0.4315454363822937
step: 150, loss: 0.544673502445221
step: 160, loss: 0.3419196903705597
step: 170, loss: 0.4826936721801758
step: 180, loss: 0.5300728678703308
step: 190, loss: 0.4858194887638092
step: 200, loss: 0.553891658782959
step: 210, loss: 0.48898276686668396
step: 220, loss: 0.3917602002620697
step: 230, loss: 0.40016552805900574
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.35826021432876587
step: 10, loss: 0.5827445983886719
step: 20, loss: 0.5822977423667908
step: 30, loss: 0.4928061366081238
step: 40, loss: 0.24979239702224731
step: 50, loss: 0.3401173949241638
step: 60, loss: 0.468505859375
step: 70, loss: 0.4256240129470825
step: 80, loss: 0.4527018666267395
step: 90, loss: 0.431675523519516
step: 100, loss: 0.5766148567199707
step: 110, loss: 0.425001859664917
step: 120, loss: 0.5466886162757874
step: 130, loss: 0.578067421913147
step: 140, loss: 0.48246681690216064
step: 150, loss: 0.6571431756019592
step: 160, loss: 0.6373828053474426
step: 170, loss: 0.4996449053287506
step: 180, loss: 0.5483560562133789
step: 190, loss: 0.4906216561794281
step: 200, loss: 0.4635443687438965
step: 210, loss: 0.47931140661239624
step: 220, loss: 0.47532176971435547
step: 230, loss: 0.38055846095085144
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5886766910552979
step: 10, loss: 0.4253317713737488
step: 20, loss: 0.5472833514213562
step: 30, loss: 0.5936130881309509
step: 40, loss: 0.6057833433151245
step: 50, loss: 0.585813581943512
step: 60, loss: 0.5396593809127808
step: 70, loss: 0.5401356220245361
step: 80, loss: 0.3362346291542053
step: 90, loss: 0.4857032895088196
step: 100, loss: 0.48228275775909424
step: 110, loss: 0.39648404717445374
step: 120, loss: 0.48821860551834106
step: 130, loss: 0.4426977038383484
step: 140, loss: 0.5297033190727234
step: 150, loss: 0.568228006362915

```



```

step: 160, loss: 0.4899738132953644
step: 170, loss: 0.5230871438980103
step: 180, loss: 0.4365713596343994
step: 190, loss: 0.57989102602005
step: 200, loss: 0.39041438698768616
step: 210, loss: 0.48530513048171997
step: 220, loss: 0.304974764585495
step: 230, loss: 0.4838327467441559
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39531978964805603
step: 10, loss: 0.4817093312740326
step: 20, loss: 0.6221597790718079
step: 30, loss: 0.3473460376262665
step: 40, loss: 0.5975106358528137
step: 50, loss: 0.6299595236778259
step: 60, loss: 0.4906327426433563
step: 70, loss: 0.3750316798686981
step: 80, loss: 0.33149275183677673
step: 90, loss: 0.43886351585388184
step: 100, loss: 0.5794224143028259
step: 110, loss: 0.5718334317207336
step: 120, loss: 0.526775598526001
step: 130, loss: 0.5729429125785828
step: 140, loss: 0.4767100214958191
step: 150, loss: 0.2984195649623871
step: 160, loss: 0.5298348069190979
step: 170, loss: 0.5804843306541443
step: 180, loss: 0.5739737153053284
step: 190, loss: 0.5758187174797058
step: 200, loss: 0.3014812171459198
step: 210, loss: 0.5300965309143066
step: 220, loss: 0.5780726671218872
step: 230, loss: 0.48810553550720215
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.4293999969959259
step: 10, loss: 0.37389224767684937
step: 20, loss: 0.480271577835083
step: 30, loss: 0.6791380047798157
step: 40, loss: 0.4380270838737488
step: 50, loss: 0.4390391707420349
step: 60, loss: 0.4868130087852478
step: 70, loss: 0.44251638650894165
step: 80, loss: 0.4389912486076355
step: 90, loss: 0.5874529480934143
step: 100, loss: 0.48213380575180054
step: 110, loss: 0.6389923095703125
step: 120, loss: 0.2950808107852936
step: 130, loss: 0.6783213019371033
step: 140, loss: 0.48482745885849
step: 150, loss: 0.34792956709861755
step: 160, loss: 0.48287874460220337
step: 170, loss: 0.5342674851417542
step: 180, loss: 0.5365498065948486
step: 190, loss: 0.4468235671520233
step: 200, loss: 0.5708290934562683
step: 210, loss: 0.42998743057250977
step: 220, loss: 0.537386953830719
step: 230, loss: 0.3491886854171753
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4366171956062317
step: 10, loss: 0.5261189937591553
step: 20, loss: 0.48209887742996216
step: 30, loss: 0.5390468835830688
step: 40, loss: 0.3844020664691925
step: 50, loss: 0.49211016297340393
step: 60, loss: 0.38434892892837524
step: 70, loss: 0.4367552101612091
step: 80, loss: 0.48922988772392273
step: 90, loss: 0.43029382824897766
step: 100, loss: 0.38242384791374207
step: 110, loss: 0.5851330757141113
step: 120, loss: 0.6348317265510559
step: 130, loss: 0.4303578734397888
step: 140, loss: 0.5786835551261902
step: 150, loss: 0.5256639719009399

```

```

step: 160, loss: 0.43940216302871704
step: 170, loss: 0.34732121229171753
step: 180, loss: 0.5280331373214722
step: 190, loss: 0.3370090126991272
step: 200, loss: 0.4866493344306946
step: 210, loss: 0.383772611618042
step: 220, loss: 0.5316476821899414
step: 230, loss: 0.5833815932273865
/usr/local/lib/python3.7/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=0.3044223517312307, f1=0.3044223517312307,
best_f1=0.3044223517312307
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:15, 157.35it/s]
load_f1 = 0.3044223517312307
real_f1 = 0.3044223517312307
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, 143.00it/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 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: 0.6228115558624268
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: 0.41177400946617126
step: 20, loss: 0.3063619136810303
step: 30, loss: 0.3733733892440796
step: 40, loss: 0.28684401512145996
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 50, loss: 0.3765451908111572
step: 60, loss: 0.27708882093429565
step: 70, loss: 0.2307143658399582
step: 80, loss: 0.129097580909729
step: 90, loss: 0.335926353931427
step: 100, loss: 0.17783096432685852
step: 110, loss: 0.19151164591312408
step: 120, loss: 0.14653493463993073
step: 130, loss: 0.1415550857782364
step: 140, loss: 0.2269909530878067
step: 150, loss: 0.0943356305360794
step: 160, loss: 0.10399944335222244
step: 170, loss: 0.07325508445501328
step: 180, loss: 0.057078056037425995
step: 190, loss: 0.050355520099401474
step: 200, loss: 0.10039485991001129
step: 210, loss: 0.07578036189079285
step: 220, loss: 0.0466504767537117
step: 230, loss: 0.14221623539924622
step: 240, loss: 0.04376757889986038
step: 250, loss: 0.10255169123411179
step: 260, loss: 0.14826180040836334

```

```

step: 270, loss: 0.38736799359321594
step: 280, loss: 0.055032484233379364
step: 290, loss: 0.03234940022230148
step: 300, loss: 0.08921452611684799
step: 310, loss: 0.05294793099164963
step: 320, loss: 0.0988490879535675
step: 330, loss: 0.04854058474302292
step: 340, loss: 0.41841357946395874
step: 350, loss: 0.1350598931312561
step: 360, loss: 0.07307718694210052
step: 370, loss: 0.037202995270490646
step: 380, loss: 0.05105892941355705
step: 390, loss: 0.028524693101644516
step: 400, loss: 0.0131845623254776
step: 410, loss: 0.30186954140663147
step: 420, loss: 0.05473436787724495
step: 430, loss: 0.05414538085460663
step: 440, loss: 0.025579897686839104
step: 450, loss: 0.1175193339586258
step: 460, loss: 0.01607310213148594
step: 470, loss: 0.031581323593854904
step: 480, loss: 0.162628635764122
step: 490, loss: 0.19134429097175598
step: 500, loss: 0.007506128866225481
step: 510, loss: 0.009968338534235954
step: 520, loss: 0.05101212486624718
step: 530, loss: 0.04219941049814224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9482999534233814, f1=0.9380281690140844,
best_f1=0.9380281690140844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053488753736019135
step: 10, loss: 0.09485463052988052
step: 20, loss: 0.006238302681595087
step: 30, loss: 0.19874218106269836
step: 40, loss: 0.09148292988538742
step: 50, loss: 0.06733730435371399
step: 60, loss: 0.047489479184150696
step: 70, loss: 0.010982438921928406
step: 80, loss: 0.02701934240758419

```

```
step: 90, loss: 0.00422185892239213
step: 100, loss: 0.03835991397500038
step: 110, loss: 0.009727953001856804
step: 120, loss: 0.05637192353606224
step: 130, loss: 0.009128192439675331
step: 140, loss: 0.10132705420255661
step: 150, loss: 0.004019975196570158
step: 160, loss: 0.05184206739068031
step: 170, loss: 0.02444293536245823
step: 180, loss: 0.05493545159697533
step: 190, loss: 0.014966442249715328
step: 200, loss: 0.03455928713083267
step: 210, loss: 0.036743685603141785
step: 220, loss: 0.0021247672848403454
step: 230, loss: 0.05978582426905632
step: 240, loss: 0.076692596077919
step: 250, loss: 0.00559297576546669
step: 260, loss: 0.08457793295383453
step: 270, loss: 0.026098325848579407
step: 280, loss: 0.05317828804254532
step: 290, loss: 0.047355953603982925
step: 300, loss: 0.044456347823143005
step: 310, loss: 0.07614554464817047
step: 320, loss: 0.18773022294044495
step: 330, loss: 0.1712600290775299
step: 340, loss: 0.04362643510103226
step: 350, loss: 0.006191635970026255
step: 360, loss: 0.07725353538990021
step: 370, loss: 0.013356766663491726
step: 380, loss: 0.14825153350830078
step: 390, loss: 0.010192127898335457
step: 400, loss: 0.06649352610111237
step: 410, loss: 0.07663430273532867
step: 420, loss: 0.03976849466562271
step: 430, loss: 0.15423589944839478
step: 440, loss: 0.015652023255825043
step: 450, loss: 0.07992509007453918
step: 460, loss: 0.07977841794490814
step: 470, loss: 0.06004452332854271
step: 480, loss: 0.00800387654453516
step: 490, loss: 0.07953733205795288
step: 500, loss: 0.011213801801204681
step: 510, loss: 0.011540092527866364
step: 520, loss: 0.24998363852500916
step: 530, loss: 0.11818769574165344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9488584474885845, f1=0.9403122130394856,
best_f1=0.9403122130394856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09622732549905777
step: 10, loss: 0.039274029433727264
step: 20, loss: 0.02099345251917839
step: 30, loss: 0.03821662813425064
step: 40, loss: 0.023964792490005493
step: 50, loss: 0.016593752428889275
step: 60, loss: 0.02948787808418274
step: 70, loss: 0.018933670595288277
step: 80, loss: 0.058520156890153885
step: 90, loss: 0.06240082532167435
step: 100, loss: 0.0739438533782959
step: 110, loss: 0.11243954300880432
step: 120, loss: 0.15068162977695465
step: 130, loss: 0.015796804800629616
step: 140, loss: 0.0037238227669149637
step: 150, loss: 0.025709250941872597
step: 160, loss: 0.229063019156456
step: 170, loss: 0.003146421629935503
step: 180, loss: 0.005349443294107914
step: 190, loss: 0.0037051448598504066
step: 200, loss: 0.012959708459675312
step: 210, loss: 0.03514948859810829
step: 220, loss: 0.09557752311229706
step: 230, loss: 0.03513574227690697
step: 240, loss: 0.026240747421979904
step: 250, loss: 0.024081163108348846
step: 260, loss: 0.13277587294578552
step: 270, loss: 0.0019436993170529604
step: 280, loss: 0.002692535985261202
step: 290, loss: 0.010023219510912895
step: 300, loss: 0.09705909341573715
step: 310, loss: 0.08609972149133682
step: 320, loss: 0.032419901341199875
step: 330, loss: 0.02852408029139042
step: 340, loss: 0.04823662340641022
step: 350, loss: 0.11074833571910858
step: 360, loss: 0.11313329637050629
step: 370, loss: 0.01447615772485733
step: 380, loss: 0.04227536544203758

```



```

step: 390, loss: 0.007613025140017271
step: 400, loss: 0.15146005153656006
step: 410, loss: 0.09647336602210999
step: 420, loss: 0.009205799549818039
step: 430, loss: 0.047371186316013336
step: 440, loss: 0.20037595927715302
step: 450, loss: 0.02604079619050026
step: 460, loss: 0.07093129307031631
step: 470, loss: 0.042954109609127045
step: 480, loss: 0.2979399561882019
step: 490, loss: 0.02379811927676201
step: 500, loss: 0.017231430858373642
step: 510, loss: 0.11570664495229721
step: 520, loss: 0.019114123657345772
step: 530, loss: 0.011552362702786922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9551548774849746, 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: 0.02611473575234413
step: 10, loss: 0.0056772963143885136
step: 20, loss: 0.054530687630176544
step: 30, loss: 0.13270023465156555
step: 40, loss: 0.009518229402601719
step: 50, loss: 0.07395996898412704
step: 60, loss: 0.003592440625652671
step: 70, loss: 0.01360565796494484
step: 80, loss: 0.030632320791482925
step: 90, loss: 0.06592939049005508
step: 100, loss: 0.002983906539157033
step: 110, loss: 0.19126063585281372
step: 120, loss: 0.0011456599459052086
step: 130, loss: 0.016863590106368065
step: 140, loss: 0.0384855642914772
step: 150, loss: 0.0009032421512529254
step: 160, loss: 0.023032410070300102
step: 170, loss: 0.06199014559388161
step: 180, loss: 0.046991027891635895
step: 190, loss: 0.08838508278131485
step: 200, loss: 0.04858480021357536

```

```

step: 210, loss: 0.018887659534811974
step: 220, loss: 0.0013148869620636106
step: 230, loss: 0.011708483099937439
step: 240, loss: 0.007800032384693623
step: 250, loss: 0.18190939724445343
step: 260, loss: 0.0015188832767307758
step: 270, loss: 0.12958969175815582
step: 280, loss: 0.011321581900119781
step: 290, loss: 0.03991380333900452
step: 300, loss: 0.002812217688187957
step: 310, loss: 0.0027320869266986847
step: 320, loss: 0.14409087598323822
step: 330, loss: 0.01896297000348568
step: 340, loss: 0.0026687828358262777
step: 350, loss: 0.10623423010110855
step: 360, loss: 0.0948457270860672
step: 370, loss: 0.0027784546837210655
step: 380, loss: 0.012334248051047325
step: 390, loss: 0.001496500801295042
step: 400, loss: 0.01387381087988615
step: 410, loss: 0.0014629876241087914
step: 420, loss: 0.010491259396076202
step: 430, loss: 0.004380217753350735
step: 440, loss: 0.0025959359481930733
step: 450, loss: 0.004822833929210901
step: 460, loss: 0.012989911250770092
step: 470, loss: 0.005622394382953644
step: 480, loss: 0.01070568710565567
step: 490, loss: 0.009279214777052402
step: 500, loss: 0.07393151521682739
step: 510, loss: 0.04409683495759964
step: 520, loss: 0.011749918572604656
step: 530, loss: 0.10887152701616287
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9584870848708487, f1=0.9459584295612009,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037760655395686626
step: 10, loss: 0.0546330027282238
step: 20, loss: 0.0030169032979756594

```

step: 30, loss: 0.04043269902467728
step: 40, loss: 0.0014630047371611
step: 50, loss: 0.08978906273841858
step: 60, loss: 0.017869189381599426
step: 70, loss: 0.0007981817470863461
step: 80, loss: 0.00048403305117972195
step: 90, loss: 0.003614316461607814
step: 100, loss: 0.30659961700439453
step: 110, loss: 0.0022563764359802008
step: 120, loss: 0.14835228025913239
step: 130, loss: 0.012641724199056625
step: 140, loss: 0.0038766618818044662
step: 150, loss: 0.017590230330824852
step: 160, loss: 0.004886139649897814
step: 170, loss: 0.08547881245613098
step: 180, loss: 0.02926095388829708
step: 190, loss: 0.009955178014934063
step: 200, loss: 0.0041907597333192825
step: 210, loss: 0.0010820418829098344
step: 220, loss: 0.013628092594444752
step: 230, loss: 0.0023275120183825493
step: 240, loss: 0.0191463865339756
step: 250, loss: 0.11927463114261627
step: 260, loss: 0.0027179259341210127
step: 270, loss: 0.005210648290812969
step: 280, loss: 0.00585935590788722
step: 290, loss: 0.001124450471252203
step: 300, loss: 0.00947335921227932
step: 310, loss: 0.12636476755142212
step: 320, loss: 0.03714572638273239
step: 330, loss: 0.025377411395311356
step: 340, loss: 0.015833985060453415
step: 350, loss: 0.05989721417427063
step: 360, loss: 0.0016466114902868867
step: 370, loss: 0.0018424749141559005
step: 380, loss: 0.00032956135692074895
step: 390, loss: 0.0016058955807238817
step: 400, loss: 0.008469728752970695
step: 410, loss: 0.022665340453386307
step: 420, loss: 0.2542954981327057
step: 430, loss: 0.013427152298390865
step: 440, loss: 0.00676168967038393
step: 450, loss: 0.003754256060346961
step: 460, loss: 0.001214488293044269
step: 470, loss: 0.022358229383826256
step: 480, loss: 0.0025935436133295298
step: 490, loss: 0.006447182036936283
step: 500, loss: 0.003992118407040834

```

step: 510, loss: 0.002705225721001625
step: 520, loss: 0.039745260030031204
step: 530, loss: 0.02188541181385517
/usr/local/lib/python3.7/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=0.9530013959981387, f1=0.9441052137153593,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005320139694958925
step: 10, loss: 0.00020272994879633188
step: 20, loss: 0.0003176526806782931
step: 30, loss: 0.001202762359753251
step: 40, loss: 0.0010479082120582461
step: 50, loss: 0.0018990072421729565
step: 60, loss: 0.045868515968322754
step: 70, loss: 0.0037076789885759354
step: 80, loss: 0.0024634734727442265
step: 90, loss: 0.0013579173246398568
step: 100, loss: 0.0009872319642454386
step: 110, loss: 0.001040748436935246
step: 120, loss: 0.09007822722196579
step: 130, loss: 0.003067115554586053
step: 140, loss: 0.005901833530515432
step: 150, loss: 0.00939728133380413
step: 160, loss: 0.041047606617212296
step: 170, loss: 0.0007950259023346007
step: 180, loss: 0.008707883767783642
step: 190, loss: 0.02537328377366066
step: 200, loss: 0.004449239932000637
step: 210, loss: 0.003026033053174615
step: 220, loss: 0.004348006099462509
step: 230, loss: 0.01892729476094246
step: 240, loss: 0.009962541051208973
step: 250, loss: 0.03590204194188118
step: 260, loss: 0.000866335176397115
step: 270, loss: 0.0037770806811749935
step: 280, loss: 0.0005024512647651136
step: 290, loss: 0.014279128983616829
step: 300, loss: 0.014475447125732899
step: 310, loss: 0.03676171228289604
step: 320, loss: 0.0014799907803535461
step: 330, loss: 0.000833591097034514
step: 340, loss: 0.0007507033878937364
step: 350, loss: 0.000978947733528912
step: 360, loss: 0.01523949857801199

```

```

step: 370, loss: 0.011485740542411804
step: 380, loss: 0.0002877051883842796
step: 390, loss: 0.0013305548345670104
step: 400, loss: 0.014374717138707638
step: 410, loss: 0.0005695832660421729
step: 420, loss: 0.055433712899684906
step: 430, loss: 0.0009438138804398477
step: 440, loss: 0.0009040726581588387
step: 450, loss: 0.17441844940185547
step: 460, loss: 0.005061980802565813
step: 470, loss: 0.002999455202370882
step: 480, loss: 0.0014543758006766438
step: 490, loss: 0.01647663116455078
step: 500, loss: 0.001341692404821515
step: 510, loss: 0.16331255435943604
step: 520, loss: 0.0018498061690479517
step: 530, loss: 0.0011851744493469596
/usr/local/lib/python3.7/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=0.9526000920386564, f1=0.944932901434521,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004549640289042145
step: 10, loss: 0.0029721148312091827
step: 20, loss: 0.011684209108352661
step: 30, loss: 0.015583273023366928
step: 40, loss: 0.0012775263749063015
step: 50, loss: 0.0006068829679861665
step: 60, loss: 0.004555799067020416
step: 70, loss: 0.00024493635282851756
step: 80, loss: 0.00021313012985046953
step: 90, loss: 4.4504606194095686e-05
step: 100, loss: 0.011180522851645947
step: 110, loss: 6.168585241539404e-05
step: 120, loss: 9.521083120489493e-05
step: 130, loss: 4.65195371361915e-05
step: 140, loss: 0.0006427796324715018
step: 150, loss: 0.0033050584606826305
step: 160, loss: 0.00033842347329482436
step: 170, loss: 0.027918599545955658
step: 180, loss: 0.0008232449181377888
step: 190, loss: 0.0029157421085983515
step: 200, loss: 0.0010050254641100764
step: 210, loss: 0.0008622542954981327
step: 220, loss: 0.0004051860305480659

```

```

step: 230, loss: 0.000118572439532727
step: 240, loss: 0.001265435479581356
step: 250, loss: 0.01892280764877796
step: 260, loss: 0.0003168793336953968
step: 270, loss: 0.0009334388887509704
step: 280, loss: 0.0051439786329865456
step: 290, loss: 0.0004901010543107986
step: 300, loss: 0.017163993790745735
step: 310, loss: 0.0008847528952173889
step: 320, loss: 0.004138958640396595
step: 330, loss: 0.0001742175518302247
step: 340, loss: 0.00028197135543450713
step: 350, loss: 0.0014301121700555086
step: 360, loss: 0.035537030547857285
step: 370, loss: 0.011248264461755753
step: 380, loss: 0.0238724946975708
step: 390, loss: 0.01347803883254528
step: 400, loss: 0.004539377521723509
step: 410, loss: 0.0003769883478526026
step: 420, loss: 0.043635789304971695
step: 430, loss: 0.0013026385568082333
step: 440, loss: 0.003727828385308385
step: 450, loss: 0.01477111130952835
step: 460, loss: 0.0013772117672488093
step: 470, loss: 0.15015198290348053
step: 480, loss: 0.006719948258250952
step: 490, loss: 0.016232585534453392
step: 500, loss: 0.003182468004524708
step: 510, loss: 0.013360223732888699
step: 520, loss: 0.008716273121535778
step: 530, loss: 0.0004747125203721225
/usr/local/lib/python3.7/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=0.9568445475638051, f1=0.9490892106492292,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022418395383283496
step: 10, loss: 0.001306720427237451
step: 20, loss: 0.0010726955952122808
step: 30, loss: 0.00068196642678231
step: 40, loss: 0.00024735028273425996
step: 50, loss: 0.000297330814646557
step: 60, loss: 0.004879667889326811
step: 70, loss: 0.00029989416361786425
step: 80, loss: 0.001156330923549831

```

```
step: 90, loss: 0.0001931221631821245
step: 100, loss: 0.0004671920614782721
step: 110, loss: 0.00048692815471440554
step: 120, loss: 0.0010509913554415107
step: 130, loss: 0.0002517307293601334
step: 140, loss: 0.003316552611067891
step: 150, loss: 0.015710311010479927
step: 160, loss: 0.00022084079682826996
step: 170, loss: 0.04245280101895332
step: 180, loss: 0.00043070537503808737
step: 190, loss: 0.0066309436224401
step: 200, loss: 0.00025382605963386595
step: 210, loss: 0.04824585095047951
step: 220, loss: 0.002430528402328491
step: 230, loss: 0.1788693219423294
step: 240, loss: 0.00042203147313557565
step: 250, loss: 0.0003823190345428884
step: 260, loss: 9.283881081501022e-05
step: 270, loss: 0.0007659745751880109
step: 280, loss: 0.00019396429706830531
step: 290, loss: 0.0002355304459342733
step: 300, loss: 4.303195237298496e-05
step: 310, loss: 0.0003957302833441645
step: 320, loss: 0.00042202536133117974
step: 330, loss: 0.00014596003165934235
step: 340, loss: 0.00834820605814457
step: 350, loss: 0.0005135886603966355
step: 360, loss: 0.007185505703091621
step: 370, loss: 0.01294993981719017
step: 380, loss: 0.00042932823998853564
step: 390, loss: 0.012086618691682816
step: 400, loss: 0.0001807205844670534
step: 410, loss: 0.002156958682462573
step: 420, loss: 0.00016155025514308363
step: 430, loss: 0.007345413323491812
step: 440, loss: 0.0008857696666382253
step: 450, loss: 0.0006091045215725899
step: 460, loss: 0.011808130890130997
step: 470, loss: 0.035724930465221405
step: 480, loss: 0.020788902416825294
step: 490, loss: 0.0019019396277144551
step: 500, loss: 0.0010673231445252895
step: 510, loss: 0.0004142057441640645
step: 520, loss: 0.00036032492062076926
step: 530, loss: 8.11338031780906e-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=0.9541454377026402, f1=0.9437470943747095,  
best_f1=0.9459584295612009  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 5.7065240980591625e-05  
step: 10, loss: 0.02165469527244568  
step: 20, loss: 0.00013406768266577274  
step: 30, loss: 0.07432902604341507  
step: 40, loss: 0.0003417697735130787  
step: 50, loss: 0.0014543782453984022  
step: 60, loss: 0.002125146333128214  
step: 70, loss: 0.0007614223286509514  
step: 80, loss: 0.026533333584666252  
step: 90, loss: 0.050685979425907135  
step: 100, loss: 0.00044246992911212146  
step: 110, loss: 0.014165811240673065  
step: 120, loss: 0.004732160829007626  
step: 130, loss: 0.00047324184561148286  
step: 140, loss: 0.00011587439803406596  
step: 150, loss: 0.0020965486764907837  
step: 160, loss: 0.0003823799197562039  
step: 170, loss: 0.010370062664151192  
step: 180, loss: 0.00022890367836225778  
step: 190, loss: 8.65500551299192e-05  
step: 200, loss: 0.0017197771230712533  
step: 210, loss: 0.0013430423568934202  
step: 220, loss: 0.0009324035490863025  
step: 230, loss: 0.0003269965236540884  
step: 240, loss: 6.172397115733474e-05  
step: 250, loss: 0.003558143973350525  
step: 260, loss: 0.0011035777861252427  
step: 270, loss: 0.04046124219894409  
step: 280, loss: 0.03832586854696274  
step: 290, loss: 0.0003932946128770709  
step: 300, loss: 0.005714574363082647  
step: 310, loss: 0.04610651358962059  
step: 320, loss: 0.01457943208515644  
step: 330, loss: 0.0018668174743652344  
step: 340, loss: 0.028012679889798164  
step: 350, loss: 0.01790766976773739  
step: 360, loss: 0.00035791617119684815  
step: 370, loss: 0.00304412585683167  
step: 380, loss: 0.00048611537204124033  
step: 390, loss: 6.914480763953179e-05  
step: 400, loss: 0.009697502478957176  
step: 410, loss: 0.0011004244443029165  
step: 420, loss: 0.00029117288067936897
```



```

step: 430, loss: 0.019926149398088455
step: 440, loss: 0.0014466712018474936
step: 450, loss: 0.0027382734697312117
step: 460, loss: 0.0021771329920738935
step: 470, loss: 2.7383663109503686e-05
step: 480, loss: 5.071867781225592e-05
step: 490, loss: 0.0006726321298629045
step: 500, loss: 5.7102206483250484e-05
step: 510, loss: 8.153513772413135e-05
step: 520, loss: 0.027048546820878983
step: 530, loss: 0.12187676876783371
/usr/local/lib/python3.7/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=0.9534347625633932, f1=0.9475620975160993,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01266807783395052
step: 10, loss: 0.0008622159948572516
step: 20, loss: 0.0005785104585811496
step: 30, loss: 0.0006892187520861626
step: 40, loss: 0.0015675731701776385
step: 50, loss: 0.00013130730076227337
step: 60, loss: 0.0017128624022006989
step: 70, loss: 2.9410064598778263e-05
step: 80, loss: 3.8248126656981185e-05
step: 90, loss: 3.428784111747518e-05
step: 100, loss: 0.000176185232703574
step: 110, loss: 0.012436369433999062
step: 120, loss: 2.0242752725607716e-05
step: 130, loss: 7.125626871129498e-05
step: 140, loss: 5.495799268828705e-05
step: 150, loss: 4.0662802348379046e-05
step: 160, loss: 0.02556086704134941
step: 170, loss: 0.00016101945948321372
step: 180, loss: 0.029322968795895576
step: 190, loss: 1.7437785572838038e-05
step: 200, loss: 9.841261635301635e-05
step: 210, loss: 0.026593154296278954
step: 220, loss: 0.00017357473552692682
step: 230, loss: 3.477733480394818e-05
step: 240, loss: 8.83501343196258e-05
step: 250, loss: 0.00021095486590638757
step: 260, loss: 0.00011102305870736018
step: 270, loss: 3.094864950980991e-05
step: 280, loss: 0.044389478862285614

```

```

step: 290, loss: 0.001573151908814907
step: 300, loss: 0.0013495993334800005
step: 310, loss: 0.033039212226867676
step: 320, loss: 0.00040749041363596916
step: 330, loss: 0.001323978416621685
step: 340, loss: 4.5508943003369495e-05
step: 350, loss: 0.0005917581147514284
step: 360, loss: 2.4571330868639052e-05
step: 370, loss: 0.11863145977258682
step: 380, loss: 0.011700564064085484
step: 390, loss: 9.150130790658295e-05
step: 400, loss: 0.001109197037294507
step: 410, loss: 0.00076304777758196
step: 420, loss: 7.936178735690191e-05
step: 430, loss: 0.0001580798561917618
step: 440, loss: 0.0008971723145805299
step: 450, loss: 8.832629828248173e-05
step: 460, loss: 0.0001062387163983658
step: 470, loss: 0.00040693586925044656
step: 480, loss: 0.00021573131380137056
step: 490, loss: 0.0006450031069107354
step: 500, loss: 0.0034461715258657932
step: 510, loss: 2.2727097530150786e-05
step: 520, loss: 0.021948710083961487
step: 530, loss: 0.05467686429619789
/usr/local/lib/python3.7/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=0.9561567164179106, f1=0.95, best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036610662937164307
step: 10, loss: 0.000342479906976223
step: 20, loss: 0.0009987952653318644
step: 30, loss: 7.158317748690024e-05
step: 40, loss: 0.016235072165727615
step: 50, loss: 0.0006037780549377203
step: 60, loss: 3.580332122510299e-05
step: 70, loss: 3.109378303634003e-05
step: 80, loss: 0.0003806853201240301
step: 90, loss: 0.004863324575126171
step: 100, loss: 0.001358703593723476
step: 110, loss: 0.000169948601978831
step: 120, loss: 0.00016566830163355917
step: 130, loss: 2.8943974029971287e-05
step: 140, loss: 0.0003333264321554452
step: 150, loss: 7.944935350678861e-05

```

```

step: 160, loss: 3.609515988500789e-05
step: 170, loss: 3.0587059882236645e-05
step: 180, loss: 1.4811535947956145e-05
step: 190, loss: 0.0281740203499794
step: 200, loss: 1.7907050278154202e-05
step: 210, loss: 5.633578257402405e-05
step: 220, loss: 0.043872665613889694
step: 230, loss: 0.0015442707808688283
step: 240, loss: 0.009797913953661919
step: 250, loss: 4.094335963600315e-05
step: 260, loss: 0.006339557468891144
step: 270, loss: 0.00015821977285668254
step: 280, loss: 0.0013890627305954695
step: 290, loss: 7.216549420263618e-05
step: 300, loss: 1.7802705770009197e-05
step: 310, loss: 0.03369764983654022
step: 320, loss: 0.00025781994918361306
step: 330, loss: 1.3220842447481118e-05
step: 340, loss: 0.0008437487995252013
step: 350, loss: 2.3717997464700602e-05
step: 360, loss: 3.189027847838588e-05
step: 370, loss: 5.534094088943675e-05
step: 380, loss: 0.0013492200523614883
step: 390, loss: 0.000742260948754847
step: 400, loss: 9.261318336939439e-05
step: 410, loss: 0.0004946623230352998
step: 420, loss: 2.064125146716833e-05
step: 430, loss: 0.0008656393620185554
step: 440, loss: 5.7207191275665537e-05
step: 450, loss: 0.0005876332288607955
step: 460, loss: 3.628698686952703e-05
step: 470, loss: 0.001994363497942686
step: 480, loss: 0.0004584877460729331
step: 490, loss: 0.0003804817679338157
step: 500, loss: 0.005029379390180111
step: 510, loss: 0.00010243661381537095
step: 520, loss: 3.675301923067309e-05
step: 530, loss: 0.0009885765612125397
/usr/local/lib/python3.7/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=0.9568075117370891, f1=0.9470725995316158,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008793261367827654
step: 10, loss: 0.0002647807414177805

```

step: 20, loss: 0.0019052624702453613
step: 30, loss: 0.00010655707592377439
step: 40, loss: 0.050168927758932114
step: 50, loss: 2.289863368787337e-05
step: 60, loss: 4.370381066109985e-05
step: 70, loss: 3.9286951505346224e-05
step: 80, loss: 5.1516144594643265e-05
step: 90, loss: 8.619928121333942e-05
step: 100, loss: 0.008445503190159798
step: 110, loss: 4.677839751821011e-05
step: 120, loss: 6.792754720663652e-05
step: 130, loss: 0.00012876962136942893
step: 140, loss: 7.62290510465391e-05
step: 150, loss: 2.070455411740113e-05
step: 160, loss: 0.0006108017987571657
step: 170, loss: 5.123302253196016e-05
step: 180, loss: 2.0320952899055555e-05
step: 190, loss: 9.554006828693673e-05
step: 200, loss: 0.00047591663314960897
step: 210, loss: 0.00017909226880874485
step: 220, loss: 7.011931302258745e-05
step: 230, loss: 0.0008663787739351392
step: 240, loss: 4.270066710887477e-05
step: 250, loss: 4.514285683399066e-05
step: 260, loss: 2.4470456992276013e-05
step: 270, loss: 0.001650791266001761
step: 280, loss: 0.00017410948930773884
step: 290, loss: 0.00017689196101855487
step: 300, loss: 0.0009705022093839943
step: 310, loss: 0.0011606893967837095
step: 320, loss: 0.00044950429582968354
step: 330, loss: 0.0006925727939233184
step: 340, loss: 0.00023840917856432498
step: 350, loss: 8.883557165972888e-05
step: 360, loss: 0.00021866257884539664
step: 370, loss: 4.164870915701613e-05
step: 380, loss: 4.341582825873047e-05
step: 390, loss: 0.0007924413657747209
step: 400, loss: 0.00032396134338341653
step: 410, loss: 0.00019777311536017805
step: 420, loss: 0.0002988835913129151
step: 430, loss: 0.12024335563182831
step: 440, loss: 0.0026350494008511305
step: 450, loss: 0.003597198985517025
step: 460, loss: 0.00036692936555482447
step: 470, loss: 0.0013201397377997637
step: 480, loss: 0.0005639820592477918
step: 490, loss: 0.0007620354299433529

```
step: 500, loss: 0.0002205630880780518
step: 510, loss: 0.00194444925710558891
step: 520, loss: 0.00042184401536360383
step: 530, loss: 0.0003885014448314905
/usr/local/lib/python3.7/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=0.9558270676691729, f1=0.9472693032015067,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.857586878235452e-05
step: 10, loss: 8.547316247131675e-05
step: 20, loss: 0.00014773491420783103
step: 30, loss: 3.78506701963488e-05
step: 40, loss: 8.516048546880484e-05
step: 50, loss: 0.0006266803829930723
step: 60, loss: 0.0015646109823137522
step: 70, loss: 0.0001196570010506548
step: 80, loss: 0.0002741623029578477
step: 90, loss: 0.00041108904406428337
step: 100, loss: 2.7770594897447154e-05
step: 110, loss: 0.0005509003531187773
step: 120, loss: 3.772752461372875e-05
step: 130, loss: 0.014408999122679234
step: 140, loss: 0.000520216126460582
step: 150, loss: 0.0007461466011591256
step: 160, loss: 0.0004327371425461024
step: 170, loss: 0.0030951083172112703
step: 180, loss: 5.8899124269373715e-05
step: 190, loss: 0.0069824522361159325
step: 200, loss: 0.0001657897373661399
step: 210, loss: 4.663719664677046e-05
step: 220, loss: 0.0005084191216155887
step: 230, loss: 0.00043757326784543693
step: 240, loss: 6.0605128965107724e-05
step: 250, loss: 9.773983765626326e-05
step: 260, loss: 6.808486796217039e-05
step: 270, loss: 7.405969518003985e-05
step: 280, loss: 0.0001012322973110713
step: 290, loss: 2.0708312149508856e-05
step: 300, loss: 3.9897990063764155e-05
step: 310, loss: 4.4400458136806265e-05
step: 320, loss: 0.00031207912252284586
step: 330, loss: 6.31243601674214e-05
step: 340, loss: 0.0015289700822904706
step: 350, loss: 2.1758516595582478e-05
```

```

step: 360, loss: 0.008757282979786396
step: 370, loss: 0.00028898558230139315
step: 380, loss: 2.9915900086052716e-05
step: 390, loss: 1.6323985619237646e-05
step: 400, loss: 0.00015847415488678962
step: 410, loss: 2.9430359063553624e-05
step: 420, loss: 3.6298522900324315e-05
step: 430, loss: 3.853708767564967e-05
step: 440, loss: 1.2691890333371703e-05
step: 450, loss: 2.8931543056387454e-05
step: 460, loss: 2.405353370704688e-05
step: 470, loss: 0.0004290680808480829
step: 480, loss: 1.1980286217294633e-05
step: 490, loss: 4.086167609784752e-05
step: 500, loss: 1.3053243492322508e-05
step: 510, loss: 0.00024211473646573722
step: 520, loss: 2.334548844373785e-05
step: 530, loss: 0.0006717385840602219
/usr/local/lib/python3.7/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=0.9514018691588785, f1=0.9417840375586854,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010484707308933139
step: 10, loss: 0.000596024445258081
step: 20, loss: 2.3542699636891484e-05
step: 30, loss: 0.0003150968987029046
step: 40, loss: 4.381211692816578e-05
step: 50, loss: 4.2670650145737454e-05
step: 60, loss: 0.00013987046258989722
step: 70, loss: 2.2421765606850386e-05
step: 80, loss: 0.0003561055054888129
step: 90, loss: 1.529561632196419e-05
step: 100, loss: 1.92218940355815e-05
step: 110, loss: 7.193280180217698e-05
step: 120, loss: 7.0322843384929e-05
step: 130, loss: 0.00044889922719448805
step: 140, loss: 0.00021607796952594072
step: 150, loss: 0.0023867241106927395
step: 160, loss: 1.6286745449178852e-05
step: 170, loss: 0.0029190806671977043
step: 180, loss: 8.645354682812467e-05
step: 190, loss: 0.00014371630095411092
step: 200, loss: 0.0007984364056028426
step: 210, loss: 4.078892743564211e-05

```

```

step: 220, loss: 3.215328251826577e-05
step: 230, loss: 3.760718391276896e-05
step: 240, loss: 0.0006058872677385807
step: 250, loss: 0.0047195591032505035
step: 260, loss: 2.7171356123290025e-05
step: 270, loss: 0.0015392439672723413
step: 280, loss: 0.00030405877623707056
step: 290, loss: 1.691986653895583e-05
step: 300, loss: 2.4906172257033177e-05
step: 310, loss: 0.0005096538225188851
step: 320, loss: 0.0007357547874562442
step: 330, loss: 9.937685535987839e-05
step: 340, loss: 0.00016259841504506767
step: 350, loss: 1.648408579058014e-05
step: 360, loss: 0.0002776500768959522
step: 370, loss: 0.0012697370257228613
step: 380, loss: 0.002633497817441821
step: 390, loss: 0.00038145948201417923
step: 400, loss: 0.0008882097899913788
step: 410, loss: 4.016072489321232e-05
step: 420, loss: 3.399350316612981e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 430, loss: 0.0001541841047583148
step: 440, loss: 0.00011706135410349816
step: 450, loss: 4.135549534112215e-05
step: 460, loss: 0.025933394208550453
step: 470, loss: 1.3869110262021422e-05
step: 480, loss: 1.6778423741925508e-05
step: 490, loss: 0.0009689190192148089
step: 500, loss: 0.0007605424616485834
step: 510, loss: 0.017780259251594543
step: 520, loss: 4.668462497647852e-05
step: 530, loss: 0.00042619177838787436
/usr/local/lib/python3.7/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=0.9561975768872321, f1=0.9476635514018692,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.176904283463955e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.0006709888111799955
step: 20, loss: 0.0009536987636238337
step: 30, loss: 0.03247784078121185
step: 40, loss: 4.997356518288143e-05
step: 50, loss: 0.0005616068374365568

```

step: 60, loss: 0.00036319289938546717
step: 70, loss: 1.907672776724212e-05
step: 80, loss: 0.0007681231945753098
step: 90, loss: 0.0005504198488779366
step: 100, loss: 3.1504318030783907e-05
step: 110, loss: 0.0004501217044889927
step: 120, loss: 1.8871851352741942e-05
step: 130, loss: 7.765302871121094e-05
step: 140, loss: 4.275469109416008e-05
step: 150, loss: 1.740423613227904e-05
step: 160, loss: 2.32992424571421e-05
step: 170, loss: 2.0421130102477036e-05
step: 180, loss: 5.2920480811735615e-05
step: 190, loss: 0.00020013880566693842
step: 200, loss: 0.00011473283666418865
step: 210, loss: 0.0017074818024411798
step: 220, loss: 1.362694729323266e-05
step: 230, loss: 6.227406993275508e-05
step: 240, loss: 2.172134190914221e-05
step: 250, loss: 0.0003936279972549528
step: 260, loss: 1.0453072718519252e-05
step: 270, loss: 1.5716585039626807e-05
step: 280, loss: 2.12555405596504e-05
step: 290, loss: 8.972537762019783e-05
step: 300, loss: 1.9508896002662368e-05
step: 310, loss: 0.03118567354977131
step: 320, loss: 2.2634352717432193e-05
step: 330, loss: 1.8391396224615164e-05
step: 340, loss: 2.0298679373809136e-05
step: 350, loss: 0.0024847500026226044
step: 360, loss: 4.441268902155571e-05
step: 370, loss: 7.039311458356678e-05
step: 380, loss: 1.5374023860204034e-05
step: 390, loss: 2.6634173991624266e-05
step: 400, loss: 0.1445326805114746
step: 410, loss: 7.712441583862528e-05
step: 420, loss: 2.0238812794559635e-05
step: 430, loss: 1.0866576303669717e-05
step: 440, loss: 0.04159826040267944
step: 450, loss: 0.0004653751093428582
step: 460, loss: 0.0001263718877453357
step: 470, loss: 7.288219785550609e-05
step: 480, loss: 0.0005661759641952813
step: 490, loss: 6.857551488792524e-05
step: 500, loss: 0.0021408600732684135
step: 510, loss: 1.4632596503361128e-05
step: 520, loss: 2.4801942345220596e-05
step: 530, loss: 1.2278396752662957e-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=0.9546517064048621, f1=0.9471221338324755,
best_f1=0.9459584295612009
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 [00:31, 182.50it/s]
load_f1 = 0.9569245020842984
real_f1 = 0.9537892791127542
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:29, 148.31it/s]

```

3.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5038905143737793
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.40529975295066833
step: 20, loss: 0.509297788143158
step: 30, loss: 0.4125726521015167
step: 40, loss: 0.3221047818660736
step: 50, loss: 0.42121216654777527
step: 60, loss: 0.4954480230808258
step: 70, loss: 0.30673056840896606
step: 80, loss: 0.3784002363681793
step: 90, loss: 0.2551019489765167
step: 100, loss: 0.25518280267715454
step: 110, loss: 0.25267264246940613
step: 120, loss: 0.38981419801712036
step: 130, loss: 0.23061324656009674
step: 140, loss: 0.40567296743392944
step: 150, loss: 0.2648980915546417
step: 160, loss: 0.34913283586502075
step: 170, loss: 0.0993661880493164
step: 180, loss: 0.41497644782066345
step: 190, loss: 0.3629423975944519
step: 200, loss: 0.18414747714996338
step: 210, loss: 0.2206205278635025
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5688888888888889, f1=0.5594713656387665,
best_f1=0.5594713656387665
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18548811972141266
step: 10, loss: 0.06814566999673843
step: 20, loss: 0.34480002522468567
step: 30, loss: 0.2738359570503235
step: 40, loss: 0.3451465666294098
step: 50, loss: 0.16170355677604675
step: 60, loss: 0.23509934544563293
step: 70, loss: 0.18545937538146973
step: 80, loss: 0.18786557018756866
step: 90, loss: 0.3007381856441498
step: 100, loss: 0.3421568274497986
step: 110, loss: 0.26534196734428406
step: 120, loss: 0.08260272443294525
step: 130, loss: 0.11009004712104797
step: 140, loss: 0.07072483003139496
step: 150, loss: 0.24362322688102722
step: 160, loss: 0.070172019302845
step: 170, loss: 0.2721247375011444
step: 180, loss: 0.1149827241897583
step: 190, loss: 0.3193258047103882
step: 200, loss: 0.08934931457042694
step: 210, loss: 0.16464240849018097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.654275092936803, f1=0.6862385321100918,
best_f1=0.6862385321100918
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07632432878017426
step: 10, loss: 0.034947119653224945
step: 20, loss: 0.2608543634414673
step: 30, loss: 0.07438543438911438

```

```

step: 40, loss: 0.2849120497703552
step: 50, loss: 0.10814183205366135
step: 60, loss: 0.23246054351329803
step: 70, loss: 0.054348766803741455
step: 80, loss: 0.14964331686496735
step: 90, loss: 0.021011410281062126
step: 100, loss: 0.11906767636537552
step: 110, loss: 0.07714967429637909
step: 120, loss: 0.17966286838054657
step: 130, loss: 0.19670945405960083
step: 140, loss: 0.13935145735740662
step: 150, loss: 0.06929881870746613
step: 160, loss: 0.14048121869564056
step: 170, loss: 0.1540466547012329
step: 180, loss: 0.1385006457567215
step: 190, loss: 0.03367862105369568
step: 200, loss: 0.06985320150852203
step: 210, loss: 0.17008469998836517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6975806451612904, f1=0.6694560669456067,
best_f1=0.6694560669456067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03147191181778908
step: 10, loss: 0.08153939247131348
step: 20, loss: 0.02968817576766014
step: 30, loss: 0.06384493410587311
step: 40, loss: 0.041893355548381805
step: 50, loss: 0.10551446676254272
step: 60, loss: 0.14874525368213654
step: 70, loss: 0.02417599782347679
step: 80, loss: 0.07191350311040878
step: 90, loss: 0.09190882742404938
step: 100, loss: 0.19085997343063354
step: 110, loss: 0.3342006206512451
step: 120, loss: 0.16635023057460785
step: 130, loss: 0.31178969144821167
step: 140, loss: 0.161500945687294
step: 150, loss: 0.062243975698947906
step: 160, loss: 0.12922336161136627
step: 170, loss: 0.054902877658605576

```

```

step: 180, loss: 0.05829975754022598
step: 190, loss: 0.17880558967590332
step: 200, loss: 0.0301139447838068
step: 210, loss: 0.30774205923080444
/usr/local/lib/python3.7/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=0.6748251748251748, f1=0.7015706806282723,
best_f1=0.6694560669456067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11023513972759247
step: 10, loss: 0.010713385418057442
step: 20, loss: 0.007048224098980427
step: 30, loss: 0.020129116252064705
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.03964690491557121
step: 50, loss: 0.06275755167007446
step: 60, loss: 0.10164303332567215
step: 70, loss: 0.23238949477672577
step: 80, loss: 0.11452459543943405
step: 90, loss: 0.0538795180618763
step: 100, loss: 0.08433321118354797
step: 110, loss: 0.10089744627475739
step: 120, loss: 0.08707673102617264
step: 130, loss: 0.037296660244464874
step: 140, loss: 0.05300860479474068
step: 150, loss: 0.07674676179885864
step: 160, loss: 0.022451095283031464
step: 170, loss: 0.038238923996686935
step: 180, loss: 0.09411945194005966
step: 190, loss: 0.09690429270267487
step: 200, loss: 0.03535015881061554
step: 210, loss: 0.15940234065055847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7049504950495049, f1=0.7123287671232877,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03004973754286766

```

```

step: 10, loss: 0.0658058449625969
step: 20, loss: 0.02138117142021656
step: 30, loss: 0.09182083606719971
step: 40, loss: 0.014321891590952873
step: 50, loss: 0.02245686575770378
step: 60, loss: 0.13707104325294495
step: 70, loss: 0.07205746322870255
step: 80, loss: 0.07887259870767593
step: 90, loss: 0.10715221613645554
step: 100, loss: 0.07823911309242249
step: 110, loss: 0.10002307593822479
step: 120, loss: 0.004453730769455433
step: 130, loss: 0.008433200418949127
step: 140, loss: 0.01904170587658882
step: 150, loss: 0.02553398720920086
step: 160, loss: 0.01259889081120491
step: 170, loss: 0.020303677767515182
step: 180, loss: 0.02915787324309349
step: 190, loss: 0.020459052175283432
step: 200, loss: 0.046454377472400665
step: 210, loss: 0.0422203466296196
/usr/local/lib/python3.7/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=0.6848249027237354, f1=0.6934865900383141,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.087651327252388
step: 10, loss: 0.0030958373099565506
step: 20, loss: 0.023600826039910316
step: 30, loss: 0.11085893213748932
step: 40, loss: 0.01691882312297821
step: 50, loss: 0.03746478259563446
step: 60, loss: 0.028160039335489273
step: 70, loss: 0.03949984163045883
step: 80, loss: 0.010603799484670162
step: 90, loss: 0.1010919138789177
step: 100, loss: 0.039279136806726456
step: 110, loss: 0.06870709359645844
step: 120, loss: 0.08169154822826385
step: 130, loss: 0.0984373614192009
step: 140, loss: 0.05628633499145508
step: 150, loss: 0.013695701956748962
step: 160, loss: 0.15629760921001434
step: 170, loss: 0.04323940724134445
step: 180, loss: 0.0031434965785592794

```

```

step: 190, loss: 0.0170116163790226
step: 200, loss: 0.07334133982658386
step: 210, loss: 0.06585323810577393
/usr/local/lib/python3.7/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=0.6953271028037382, f1=0.7124773960216998,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0670437216758728
step: 10, loss: 0.003929907456040382
step: 20, loss: 0.01964692771434784
step: 30, loss: 0.05295756459236145
step: 40, loss: 0.003983489703387022
step: 50, loss: 0.0004513571038842201
step: 60, loss: 0.02424994297325611
step: 70, loss: 0.11544326692819595
step: 80, loss: 0.03900555148720741
step: 90, loss: 0.060011252760887146
step: 100, loss: 0.1810678243637085
step: 110, loss: 0.06074253097176552
step: 120, loss: 0.08950259536504745
step: 130, loss: 0.00184845388866961
step: 140, loss: 0.08324374258518219
step: 150, loss: 0.11549372971057892
step: 160, loss: 0.12310164421796799
step: 170, loss: 0.04250062257051468
step: 180, loss: 0.021122798323631287
step: 190, loss: 0.008371710777282715
step: 200, loss: 0.058124762028455734
step: 210, loss: 0.048916056752204895
/usr/local/lib/python3.7/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=0.6779661016949152, f1=0.708029197080292,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04016494378447533
step: 10, loss: 0.1991264671087265
step: 20, loss: 0.020291768014431
step: 30, loss: 0.020574916154146194
step: 40, loss: 0.06302753835916519
step: 50, loss: 0.03855658695101738
step: 60, loss: 0.007623148150742054

```



```

step: 70, loss: 0.022834863513708115
step: 80, loss: 0.04156060144305229
step: 90, loss: 0.039947718381881714
step: 100, loss: 0.028761323541402817
step: 110, loss: 0.09285224974155426
step: 120, loss: 0.11307373642921448
step: 130, loss: 0.025837427005171776
step: 140, loss: 0.0838431864976883
step: 150, loss: 0.0891423374414444
step: 160, loss: 0.06842775642871857
step: 170, loss: 0.05877652019262314
step: 180, loss: 0.06527189165353775
step: 190, loss: 0.026058653369545937
step: 200, loss: 0.0035676464904099703
step: 210, loss: 0.12596476078033447
/usr/local/lib/python3.7/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=0.6956521739130436, f1=0.72265625, best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016158390790224075
step: 10, loss: 0.007063223980367184
step: 20, loss: 0.0022756957914680243
step: 30, loss: 0.009772657416760921
step: 40, loss: 0.02759419195353985
step: 50, loss: 0.009187587536871433
step: 60, loss: 0.0004273864906281233
step: 70, loss: 0.008932726457715034
step: 80, loss: 0.02162211760878563
step: 90, loss: 0.07651892304420471
step: 100, loss: 0.056955914944410324
step: 110, loss: 0.009226938709616661
step: 120, loss: 0.06104740500450134
step: 130, loss: 0.02690974995493889
step: 140, loss: 0.0044930437579751015
step: 150, loss: 0.06467745453119278
step: 160, loss: 0.03210155665874481
step: 170, loss: 0.05161861702799797
step: 180, loss: 0.00344507722184062
step: 190, loss: 0.020673204213380814
step: 200, loss: 0.03103143908083439
step: 210, loss: 0.0006598153850063682
/usr/local/lib/python3.7/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=0.6816326530612244, f1=0.7191919191919192,

```

```

best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05738205462694168
step: 10, loss: 0.0015085951890796423
step: 20, loss: 0.010376346297562122
step: 30, loss: 0.05045417696237564
step: 40, loss: 0.0006285686395131052
step: 50, loss: 0.0021087436471134424
step: 60, loss: 0.002405912149697542
step: 70, loss: 0.0012380023254081607
step: 80, loss: 0.0007038123439997435
step: 90, loss: 0.2322007417678833
step: 100, loss: 0.0011376680340617895
step: 110, loss: 0.0033318998757749796
step: 120, loss: 0.018890708684921265
step: 130, loss: 0.006598141510039568
step: 140, loss: 0.12465184926986694
step: 150, loss: 0.0016293303342536092
step: 160, loss: 0.0004435603623278439
step: 170, loss: 0.0036728482227772474
step: 180, loss: 0.0012059451546519995
step: 190, loss: 0.010961731895804405
step: 200, loss: 0.009893679060041904
step: 210, loss: 0.0010355561971664429
/usr/local/lib/python3.7/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=0.679324894514768, f1=0.7098121085594988,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011099366238340735
step: 10, loss: 0.003548320382833481
step: 20, loss: 0.008182110264897346
step: 30, loss: 0.0006687328568659723
step: 40, loss: 0.0003827592299785465
step: 50, loss: 0.004630741663277149
step: 60, loss: 3.200692663085647e-05
step: 70, loss: 0.04060141742229462
step: 80, loss: 0.01849759928882122
step: 90, loss: 0.010808390565216541
step: 100, loss: 4.109172732569277e-05
step: 110, loss: 0.00973151158541441
step: 120, loss: 0.0006604587542824447
step: 130, loss: 0.0030014850199222565

```

```

step: 140, loss: 0.02518920600414276
step: 150, loss: 4.106263804715127e-05
step: 160, loss: 0.01042639184743166
step: 170, loss: 0.11889650672674179
step: 180, loss: 0.01665399968624115
step: 190, loss: 0.0005629973602481186
step: 200, loss: 0.004931572824716568
step: 210, loss: 0.0009771949844434857
/usr/local/lib/python3.7/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=0.68762278978389, f1=0.7356321839080459,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03309272229671478
step: 10, loss: 0.024847740307450294
step: 20, loss: 0.00026289100060239434
step: 30, loss: 0.0020207944326102734
step: 40, loss: 0.02544371783733368
step: 50, loss: 0.2264793962240219
step: 60, loss: 0.0004482257354538888
step: 70, loss: 0.000554910977371037
step: 80, loss: 0.0480453297495842
step: 90, loss: 0.022016363218426704
step: 100, loss: 0.0767861157655716
step: 110, loss: 0.05414280667901039
step: 120, loss: 0.005031547509133816
step: 130, loss: 0.010265364311635494
step: 140, loss: 0.006804191041737795
step: 150, loss: 0.0009135527652688324
step: 160, loss: 0.0416538380086422
step: 170, loss: 0.0014130747877061367
step: 180, loss: 0.0012447093613445759
step: 190, loss: 0.015295194461941719
step: 200, loss: 0.041605059057474136
step: 210, loss: 0.003227738430723548
/usr/local/lib/python3.7/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=0.6883910386965376, f1=0.7301587301587301,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013218338135629892
step: 10, loss: 0.020431429147720337

```

```

step: 20, loss: 0.0007863588980399072
step: 30, loss: 0.00655131321400404
step: 40, loss: 0.0206000953912735
step: 50, loss: 0.002163640921935439
step: 60, loss: 0.0412624329328537
step: 70, loss: 0.00034005267661996186
step: 80, loss: 0.022521046921610832
step: 90, loss: 0.0003147358074784279
step: 100, loss: 0.055692996829748154
step: 110, loss: 0.005905741825699806
step: 120, loss: 0.003757880302146077
step: 130, loss: 0.013783814385533333
step: 140, loss: 0.004946720786392689
step: 150, loss: 0.0017996694659814239
step: 160, loss: 0.0005276564625091851
step: 170, loss: 0.037063367664813995
step: 180, loss: 0.010446186177432537
step: 190, loss: 0.050929922610521317
step: 200, loss: 0.07370322197675705
step: 210, loss: 0.0035630231723189354
/usr/local/lib/python3.7/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=0.6895161290322581, f1=0.7312252964426877,
best_f1=0.7123287671232877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001999960746616125
step: 10, loss: 0.0001764945627655834
step: 20, loss: 0.014643649570643902
step: 30, loss: 0.014262018725275993
step: 40, loss: 0.0016967483097687364
step: 50, loss: 0.0001922838855534792
step: 60, loss: 0.007314937189221382
step: 70, loss: 0.02966899238526821
step: 80, loss: 0.02115633338689804
step: 90, loss: 0.0002911231422331184
step: 100, loss: 0.000211995531572029
step: 110, loss: 0.210957333445549
step: 120, loss: 0.00043743415153585374
step: 130, loss: 0.007937050424516201
step: 140, loss: 0.010189315304160118
step: 150, loss: 0.0047197178937494755
step: 160, loss: 0.0032612846698611975
step: 170, loss: 0.01417555008083582
step: 180, loss: 0.00017887511057779193
step: 190, loss: 0.0006993310526013374

```

```

step: 200, loss: 0.00018658320186659694
step: 210, loss: 0.02488936297595501
/usr/local/lib/python3.7/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=0.6878980891719746, f1=0.7208333333333334,
best_f1=0.7123287671232877
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:09, 236.15it/s]
load_f1 = 0.7035573122529645
real_f1 = 0.6926229508196722
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, 141.10it/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: 0.4693189859390259
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.45733627676963806
step: 20, loss: 0.28210216760635376
step: 30, loss: 0.38483789563179016
step: 40, loss: 0.2763957679271698
step: 50, loss: 0.3511425256729126
step: 60, loss: 0.459878146648407
step: 70, loss: 0.523112952709198
step: 80, loss: 0.1627463847398758
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.29999348521232605
step: 100, loss: 0.4158961772918701
step: 110, loss: 0.20317895710468292
step: 120, loss: 0.2963840663433075
step: 130, loss: 0.3265044093132019
step: 140, loss: 0.14044441282749176
step: 150, loss: 0.1995699554681778
step: 160, loss: 0.19750520586967468
step: 170, loss: 0.45320925116539
step: 180, loss: 0.1445576697587967
step: 190, loss: 0.16868233680725098
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5847665847665847, f1=0.5790754257907542,
best_f1=0.5790754257907542
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23162490129470825
step: 10, loss: 0.10367385298013687
step: 20, loss: 0.786734402179718
step: 30, loss: 0.27749937772750854
step: 40, loss: 0.3549702763557434
step: 50, loss: 0.3778034746646881
step: 60, loss: 0.13752318918704987
step: 70, loss: 0.04647861421108246
step: 80, loss: 0.10720362514257431
step: 90, loss: 0.21435073018074036
step: 100, loss: 0.07290510833263397
step: 110, loss: 0.0847998559474945
step: 120, loss: 0.14492584764957428
step: 130, loss: 0.23901227116584778
step: 140, loss: 0.2243114858865738
step: 150, loss: 0.25217723846435547
step: 160, loss: 0.16044865548610687
step: 170, loss: 0.03756159171462059
step: 180, loss: 0.062261827290058136
step: 190, loss: 0.29166877269744873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7250608272506084, f1=0.6775700934579438,
best_f1=0.6775700934579438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.154838427901268
step: 10, loss: 0.2731775641441345
step: 20, loss: 0.16589237749576569
step: 30, loss: 0.07352736592292786
step: 40, loss: 0.07823061943054199
step: 50, loss: 0.11469665169715881
step: 60, loss: 0.015295620076358318

```



```

step: 70, loss: 0.18304039537906647
step: 80, loss: 0.13379904627799988
step: 90, loss: 0.17015524208545685
step: 100, loss: 0.1524619162082672
step: 110, loss: 0.4065856635570526
step: 120, loss: 0.09072064608335495
step: 130, loss: 0.02445446141064167
step: 140, loss: 0.03262539952993393
step: 150, loss: 0.2817055881023407
step: 160, loss: 0.07026410102844238
step: 170, loss: 0.1828324943780899
step: 180, loss: 0.10448060929775238
step: 190, loss: 0.03456133231520653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.8062015503875969, f1=0.8062015503875969,
best_f1=0.8062015503875969
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020038357004523277
step: 10, loss: 0.016826247796416283
step: 20, loss: 0.11475078761577606
step: 30, loss: 0.00790583249181509
step: 40, loss: 0.19194985926151276
step: 50, loss: 0.033765871077775955
step: 60, loss: 0.05377373471856117
step: 70, loss: 0.052074138075113297
step: 80, loss: 0.06603898853063583
step: 90, loss: 0.024127913638949394
step: 100, loss: 0.13525746762752533
step: 110, loss: 0.24796311557292938
step: 120, loss: 0.015582535415887833
step: 130, loss: 0.021106252446770668
step: 140, loss: 0.2092365324497223
step: 150, loss: 0.013184914365410805
step: 160, loss: 0.014989827759563923
step: 170, loss: 0.03982580825686455
step: 180, loss: 0.018692009150981903
step: 190, loss: 0.007543968968093395
/usr/local/lib/python3.7/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=0.7755102040816326, f1=0.78698224852071,
best_f1=0.8062015503875969
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2488364577293396
step: 10, loss: 0.053518954664468765
step: 20, loss: 0.008846618235111237
step: 30, loss: 0.021780237555503845
step: 40, loss: 0.038088444620370865
step: 50, loss: 0.01938686892390251
step: 60, loss: 0.003902496537193656
step: 70, loss: 0.04920892417430878
step: 80, loss: 0.006932667922228575
step: 90, loss: 0.030261728912591934
step: 100, loss: 0.006235482171177864
step: 110, loss: 0.1765495240688324
step: 120, loss: 0.10027091950178146
step: 130, loss: 0.31951919198036194
step: 140, loss: 0.0409020259976387
step: 150, loss: 0.058480314910411835
step: 160, loss: 0.012793687172234058
step: 170, loss: 0.2110299915075302
step: 180, loss: 0.01002564188092947
step: 190, loss: 0.0078077735379338264
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8235294117647058, f1=0.8368421052631578,
best_f1=0.8368421052631578
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20081856846809387
step: 10, loss: 0.09424681216478348
step: 20, loss: 0.04791039228439331
step: 30, loss: 0.02866315469145775
step: 40, loss: 0.08399743586778641
step: 50, loss: 0.015454920940101147
step: 60, loss: 0.007712752558290958
step: 70, loss: 0.130723774433136
step: 80, loss: 0.03836420550942421
step: 90, loss: 0.04935717210173607
step: 100, loss: 0.023793620988726616

```

```

step: 110, loss: 0.01889512501657009
step: 120, loss: 0.027366172522306442
step: 130, loss: 0.2357591986656189
step: 140, loss: 0.0700833722949028
step: 150, loss: 0.06729679554700851
step: 160, loss: 0.21504917740821838
step: 170, loss: 0.1778879314661026
step: 180, loss: 0.04382218420505524
step: 190, loss: 0.01280608307570219
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8241758241758241, f1=0.8174386920980926,
best_f1=0.8174386920980926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00820514652878046
step: 10, loss: 0.009467409923672676
step: 20, loss: 0.004369889851659536
step: 30, loss: 0.0550985224545002
step: 40, loss: 0.0038379505276679993
step: 50, loss: 0.0016726208850741386
step: 60, loss: 0.15633119642734528
step: 70, loss: 0.006906955968588591
step: 80, loss: 0.002873785560950637
step: 90, loss: 0.002035269048064947
step: 100, loss: 0.26848796010017395
step: 110, loss: 0.12933368980884552
step: 120, loss: 0.0031875385902822018
step: 130, loss: 0.06791937351226807
step: 140, loss: 0.011287637986242771
step: 150, loss: 0.039950571954250336
step: 160, loss: 0.09781645983457565
step: 170, loss: 0.00999446865171194
step: 180, loss: 0.0022078517358750105
step: 190, loss: 0.012064797803759575
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8301886792452831, f1=0.8346883468834688,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04459430277347565
step: 10, loss: 0.052154164761304855
step: 20, loss: 0.004127774853259325
step: 30, loss: 0.15573972463607788
step: 40, loss: 0.015840968117117882
step: 50, loss: 0.02107279933989048
step: 60, loss: 0.04310765117406845
step: 70, loss: 0.0013475543819367886
step: 80, loss: 0.05841093137860298
step: 90, loss: 0.015281662344932556
step: 100, loss: 0.013948176987469196
step: 110, loss: 0.050135135650634766
step: 120, loss: 0.02123415656387806
step: 130, loss: 0.0028725704178214073
step: 140, loss: 0.0037297983653843403
step: 150, loss: 0.2204710692167282
step: 160, loss: 0.03893837332725525
step: 170, loss: 0.0011541158892214298
step: 180, loss: 0.003383539617061615
step: 190, loss: 0.006281441543251276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8351063829787233, f1=0.8455284552845528,
best_f1=0.8455284552845528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029794557485729456
step: 10, loss: 0.07719513028860092
step: 20, loss: 0.001921502873301506
step: 30, loss: 0.0017415095353499055
step: 40, loss: 0.02726176753640175
step: 50, loss: 0.082124263048172
step: 60, loss: 0.055364951491355896
step: 70, loss: 0.005338670220226049
step: 80, loss: 0.00525673059746623
step: 90, loss: 0.01239362545311451
step: 100, loss: 0.0026249594520777464

```

```

step: 110, loss: 0.07877699285745621
step: 120, loss: 0.08551187813282013
step: 130, loss: 0.004556847270578146
step: 140, loss: 0.05399392172694206
step: 150, loss: 0.0041556283831596375
step: 160, loss: 0.002760161878541112
step: 170, loss: 0.0066190543584525585
step: 180, loss: 0.005825390573590994
step: 190, loss: 0.00047442771028727293
/usr/local/lib/python3.7/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=0.8010752688172043, f1=0.8156424581005588,
best_f1=0.8455284552845528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001302504213526845
step: 10, loss: 0.011697288602590561
step: 20, loss: 0.0035331666003912687
step: 30, loss: 0.0021325484849512577
step: 40, loss: 0.0017258775187656283
step: 50, loss: 0.09535378962755203
step: 60, loss: 0.0015697962371632457
step: 70, loss: 0.10822185128927231
step: 80, loss: 0.0029653022065758705
step: 90, loss: 0.0005876655923202634
step: 100, loss: 0.005283759906888008
step: 110, loss: 0.0023942156694829464
step: 120, loss: 0.0010763616301119328
step: 130, loss: 0.0017859230283647776
step: 140, loss: 0.0020554382354021072
step: 150, loss: 0.00653440086171031
step: 160, loss: 0.0019014016725122929
step: 170, loss: 0.000782924413215369
step: 180, loss: 0.030047887936234474
step: 190, loss: 0.0012749811867251992
/usr/local/lib/python3.7/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=0.8320413436692506, f1=0.84375, best_f1=0.8455284552845528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00560813769698143
step: 10, loss: 0.0006196385365910828
step: 20, loss: 0.0008044306305237114
step: 30, loss: 0.000613063049968332

```

```

step: 40, loss: 0.0011066903825849295
step: 50, loss: 0.0011197415878996253
step: 60, loss: 0.001339406124316156
step: 70, loss: 0.0008936923113651574
step: 80, loss: 0.0993209183216095
step: 90, loss: 0.0006498206639662385
step: 100, loss: 0.0025477269664406776
step: 110, loss: 0.19365599751472473
step: 120, loss: 0.00846879743039608
step: 130, loss: 0.002872287994250655
step: 140, loss: 0.0013245090376585722
step: 150, loss: 0.0015050304355099797
step: 160, loss: 0.000823484908323735
step: 170, loss: 0.00285001820884645
step: 180, loss: 0.004655797965824604
step: 190, loss: 0.0015326490392908454
/usr/local/lib/python3.7/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=0.8328767123287673, f1=0.8555858310626703,
best_f1=0.8455284552845528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022875904105603695
step: 10, loss: 0.02546132169663906
step: 20, loss: 0.002694430062547326
step: 30, loss: 0.0025984200183302164
step: 40, loss: 0.0007554035983048379
step: 50, loss: 0.0018238051561638713
step: 60, loss: 0.0413547046482563
step: 70, loss: 0.0011829857248812914
step: 80, loss: 0.0012198478216305375
step: 90, loss: 0.0007899911142885685
step: 100, loss: 0.0015140067553147674
step: 110, loss: 0.019282141700387
step: 120, loss: 0.0018818480893969536
step: 130, loss: 0.0006591361598111689
step: 140, loss: 0.008581875823438168
step: 150, loss: 0.0011656658025458455
step: 160, loss: 0.03148072212934494
step: 170, loss: 0.011592299677431583
step: 180, loss: 0.0004944551037624478
step: 190, loss: 0.0004697772965300828
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.8429752066115702, f1=0.861878453038674,
best_f1=0.861878453038674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01993701234459877
step: 10, loss: 0.0004963864339515567
step: 20, loss: 0.004529340658336878
step: 30, loss: 0.12857049703598022
step: 40, loss: 0.00206119567155838
step: 50, loss: 0.0008110112976282835
step: 60, loss: 0.08927423506975174
step: 70, loss: 0.0008776194299571216
step: 80, loss: 0.0012170759728178382
step: 90, loss: 0.0017976736417040229
step: 100, loss: 0.0006804477307014167
step: 110, loss: 0.0011096122907474637
step: 120, loss: 0.001064011245034635
step: 130, loss: 0.0011847538407891989
step: 140, loss: 0.001152328448370099
step: 150, loss: 0.0007146669086068869
step: 160, loss: 0.0017296494916081429
step: 170, loss: 0.0018132164841517806
step: 180, loss: 0.0006916641723364592
step: 190, loss: 0.20744919776916504
/usr/local/lib/python3.7/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=0.8300835654596099, f1=0.8295454545454546,
best_f1=0.861878453038674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004438603296875954
step: 10, loss: 0.0012644616654142737
step: 20, loss: 0.0006593811558559537
step: 30, loss: 0.0013492999132722616
step: 40, loss: 0.004765267018228769
step: 50, loss: 0.0006263894611038268
step: 60, loss: 0.0010599056258797646
step: 70, loss: 0.0005942791467532516
step: 80, loss: 0.0007122571114450693
step: 90, loss: 0.00045598347787745297
step: 100, loss: 0.0005208299262449145
step: 110, loss: 0.0007819808088243008
```

```

step: 120, loss: 0.0011677165748551488
step: 130, loss: 0.00041200683335773647
step: 140, loss: 0.0006169407279230654
step: 150, loss: 0.001172911492176354
step: 160, loss: 0.0010208750609308481
step: 170, loss: 0.0005045676953159273
step: 180, loss: 0.0008061882108449936
step: 190, loss: 0.0024490300565958023
/usr/local/lib/python3.7/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=0.8359788359788362, f1=0.8541114058355438,
best_f1=0.861878453038674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021121471654623747
step: 10, loss: 0.008506111800670624
step: 20, loss: 0.0034982203505933285
step: 30, loss: 0.0003200723440386355
step: 40, loss: 0.0024339305236935616
step: 50, loss: 0.00041955732740461826
step: 60, loss: 0.0004873946018051356
step: 70, loss: 0.0025453127454966307
step: 80, loss: 0.0005497372476384044
step: 90, loss: 0.0016015886794775724
step: 100, loss: 0.0006739968666806817
step: 110, loss: 0.0006004260503686965
step: 120, loss: 0.0005311177228577435
step: 130, loss: 0.0012965830974280834
step: 140, loss: 0.0009118241141550243
step: 150, loss: 0.00037732720375061035
step: 160, loss: 0.20498408377170563
step: 170, loss: 0.0012453995877876878
step: 180, loss: 0.0004511510778684169
step: 190, loss: 0.0010876004816964269
/usr/local/lib/python3.7/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=0.8401084010840109, f1=0.8455284552845528,
best_f1=0.861878453038674
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 [00:12, 161.48it/s]
load_f1 = 0.8105263157894735
real_f1 = 0.7989690721649483
733it [00:00, 3535.86it/s]/usr/local/lib/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, 140.92it/s]

```

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

```

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

```

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

step: 0, loss: 0.5262883305549622

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

step: 10, loss: 0.47739142179489136

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

step: 20, loss: 0.29900026321411133

step: 30, loss: 0.37376248836517334

step: 40, loss: 0.5116761922836304

step: 50, loss: 0.3387787938117981

step: 60, loss: 0.5311005711555481

step: 70, loss: 0.2986036241054535

step: 80, loss: 0.23588460683822632

step: 90, loss: 0.21884040534496307

step: 100, loss: 0.15152060985565186

step: 110, loss: 0.3895699977874756

step: 120, loss: 0.2644915282726288

step: 130, loss: 0.3051247298717499

step: 140, loss: 0.3704773187637329

step: 150, loss: 0.3241327106952667

step: 160, loss: 0.3937518298625946

step: 170, loss: 0.3268813490867615

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

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

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

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=0.19543023821098687, f1=0.19607843137254902,

best_f1=0.19607843137254902

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

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

step: 0, loss: 0.3168639838695526

step: 10, loss: 0.46604809165000916

step: 20, loss: 0.30765268206596375

step: 30, loss: 0.33476394414901733

step: 40, loss: 0.10692386329174042

step: 50, loss: 0.42374417185783386

```

step: 60, loss: 0.1944584995508194
step: 70, loss: 0.5086171627044678
step: 80, loss: 0.23234763741493225
step: 90, loss: 0.23639054596424103
step: 100, loss: 0.4999188184738159
step: 110, loss: 0.2616639733314514
step: 120, loss: 0.2370585948228836
step: 130, loss: 0.5429763793945312
step: 140, loss: 0.5189732313156128
step: 150, loss: 0.4198436141014099
step: 160, loss: 0.4274258017539978
step: 170, loss: 0.37036067247390747
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.23519599666388658, f1=0.2404274265360641,
best_f1=0.2404274265360641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5847660303115845
step: 10, loss: 0.3425147533416748
step: 20, loss: 0.2288687825202942
step: 30, loss: 0.2148863524198532
step: 40, loss: 0.3385932445526123
step: 50, loss: 0.4696304202079773
step: 60, loss: 0.22061903774738312
step: 70, loss: 0.19148755073547363
step: 80, loss: 0.13828341662883759
step: 90, loss: 0.4963496923446655
step: 100, loss: 0.0948626846075058
step: 110, loss: 0.05283878371119499
step: 120, loss: 0.20364171266555786
step: 130, loss: 0.2644336521625519
step: 140, loss: 0.27332642674446106
step: 150, loss: 0.2245042324066162
step: 160, loss: 0.13061794638633728
step: 170, loss: 0.09417961537837982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.794188861985472, f1=0.8018867924528301,
best_f1=0.8018867924528301
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18183907866477966
step: 10, loss: 0.12408808618783951
step: 20, loss: 0.06680990010499954
step: 30, loss: 0.13861556351184845
step: 40, loss: 0.10629978775978088
step: 50, loss: 0.07294357568025589
step: 60, loss: 0.226923868060112
step: 70, loss: 0.021504977717995644
step: 80, loss: 0.19814841449260712
step: 90, loss: 0.2597726583480835
step: 100, loss: 0.2634366750717163
step: 110, loss: 0.26453712582588196
step: 120, loss: 0.17506392300128937
step: 130, loss: 0.08664706349372864
step: 140, loss: 0.11796040087938309
step: 150, loss: 0.18804162740707397
step: 160, loss: 0.0856073871254921
step: 170, loss: 0.052272845059633255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8137254901960784, f1=0.8564705882352941,
best_f1=0.8564705882352941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.080935999751091
step: 10, loss: 0.035914815962314606
step: 20, loss: 0.2632225453853607
step: 30, loss: 0.04585853964090347
step: 40, loss: 0.08708437532186508
step: 50, loss: 0.054239243268966675
step: 60, loss: 0.09311074018478394
step: 70, loss: 0.09277959913015366
step: 80, loss: 0.0007609685999341309
step: 90, loss: 0.014411069452762604
step: 100, loss: 0.19977013766765594
step: 110, loss: 0.11086472123861313

```

```

step: 120, loss: 0.03264039382338524
step: 130, loss: 0.00738114258274436
step: 140, loss: 0.13787782192230225
step: 150, loss: 0.03401361405849457
step: 160, loss: 0.024919182062149048
step: 170, loss: 0.019716687500476837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8401084010840109, f1=0.8762886597938145,
best_f1=0.8762886597938145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07085064798593521
step: 10, loss: 0.034000858664512634
step: 20, loss: 0.11479254812002182
step: 30, loss: 0.010976001620292664
step: 40, loss: 0.028534453362226486
step: 50, loss: 0.0026677337009459734
step: 60, loss: 0.02796509489417076
step: 70, loss: 0.020959841087460518
step: 80, loss: 0.011844838969409466
step: 90, loss: 0.06458871811628342
step: 100, loss: 0.010099461302161217
step: 110, loss: 0.043870434165000916
step: 120, loss: 0.028051026165485382
step: 130, loss: 0.015667524188756943
step: 140, loss: 0.13298317790031433
step: 150, loss: 0.021654892712831497
step: 160, loss: 0.08331653475761414
step: 170, loss: 0.03245067596435547
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8523809523809525, f1=0.8584269662921348,
best_f1=0.8584269662921348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.036353014409542084
step: 10, loss: 0.005957081913948059
step: 20, loss: 0.003401684807613492
step: 30, loss: 0.0049848430790007114
step: 40, loss: 0.0009165705414488912
step: 50, loss: 0.048667434602975845
step: 60, loss: 0.012817258015275002
step: 70, loss: 0.003108395729213953
step: 80, loss: 0.041509270668029785
step: 90, loss: 0.03248889371752739
step: 100, loss: 0.0010898777982220054
step: 110, loss: 0.029316887259483337
step: 120, loss: 0.002090521389618516
step: 130, loss: 0.0009972702246159315
step: 140, loss: 0.015417328104376793
step: 150, loss: 0.11130796372890472
step: 160, loss: 0.01935208961367607
step: 170, loss: 0.024068670347332954
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8585607940446649, f1=0.8826291079812206,
best_f1=0.8826291079812206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04735400900244713
step: 10, loss: 0.017030011862516403
step: 20, loss: 0.022857245057821274
step: 30, loss: 0.0006151317502371967
step: 40, loss: 0.0009606919484212995
step: 50, loss: 0.003456605365499854
step: 60, loss: 0.0019831047393381596
step: 70, loss: 0.05602903291583061
step: 80, loss: 0.0021862743888050318
step: 90, loss: 0.05245217680931091
step: 100, loss: 0.00038511876482516527
step: 110, loss: 0.10867828875780106
step: 120, loss: 0.11758051812648773
step: 130, loss: 0.01792108453810215
step: 140, loss: 0.02867468260228634
step: 150, loss: 0.0003586544771678746
step: 160, loss: 0.0005365958204492927
step: 170, loss: 0.10695020854473114

```

```

/usr/local/lib/python3.7/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=0.8584686774941995, f1=0.8616780045351473,
best_f1=0.8826291079812206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01769915223121643
step: 10, loss: 0.042622487992048264
step: 20, loss: 0.007659326307475567
step: 30, loss: 0.008764689788222313
step: 40, loss: 0.028535693883895874
step: 50, loss: 0.00048046716256067157
step: 60, loss: 0.03185581788420677
step: 70, loss: 0.014433356001973152
step: 80, loss: 0.004195022862404585
step: 90, loss: 0.09017492830753326
step: 100, loss: 0.010375752113759518
step: 110, loss: 0.016027674078941345
step: 120, loss: 0.013133344240486622
step: 130, loss: 0.0003575048758648336
step: 140, loss: 0.00023398247139994055
step: 150, loss: 0.11631987988948822
step: 160, loss: 0.0005911399493925273
step: 170, loss: 0.01438503060489893
/usr/local/lib/python3.7/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=0.8443271767810026, f1=0.891089108910891,
best_f1=0.8826291079812206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022283976431936026
step: 10, loss: 0.0122825987637043
step: 20, loss: 0.07740499079227448
step: 30, loss: 0.2131456732749939
step: 40, loss: 0.012873256579041481
step: 50, loss: 0.015358075499534607
step: 60, loss: 0.032948657870292664
step: 70, loss: 0.026095043867826462
step: 80, loss: 0.020956125110387802
step: 90, loss: 0.005637601483613253
step: 100, loss: 0.002583675552159548
step: 110, loss: 0.017845407128334045
step: 120, loss: 0.0022689062170684338
step: 130, loss: 0.0006553717539645731

```



```

step: 140, loss: 0.0014982169959694147
step: 150, loss: 0.0008762505021877587
step: 160, loss: 0.000731658423319459
step: 170, loss: 0.02078431099653244
/usr/local/lib/python3.7/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=0.8585365853658536, f1=0.8711943793911007,
best_f1=0.8826291079812206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019971642177551985
step: 10, loss: 0.0002946988388430327
step: 20, loss: 0.0002860747335944325
step: 30, loss: 0.009401694871485233
step: 40, loss: 0.00019029354734811932
step: 50, loss: 0.007772777229547501
step: 60, loss: 0.036281611770391464
step: 70, loss: 0.0006710346206091344
step: 80, loss: 0.030049636960029602
step: 90, loss: 0.011863533407449722
step: 100, loss: 0.11993089318275452
step: 110, loss: 0.10870593041181564
step: 120, loss: 0.1654660552740097
step: 130, loss: 0.03373316302895546
step: 140, loss: 0.0011517539387568831
step: 150, loss: 0.0015260532964020967
step: 160, loss: 0.08149108290672302
step: 170, loss: 0.003682331182062626
/usr/local/lib/python3.7/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=0.8535980148883374, f1=0.8663594470046084,
best_f1=0.8826291079812206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008369560819119215
step: 10, loss: 0.0009602416539564729
step: 20, loss: 0.0004347955109551549
step: 30, loss: 0.051175929605960846
step: 40, loss: 0.0028140449430793524
step: 50, loss: 0.0007328959181904793
step: 60, loss: 0.006017169915139675
step: 70, loss: 0.00012202237121528015
step: 80, loss: 0.0059117707423865795
step: 90, loss: 0.019833993166685104

```

```

step: 100, loss: 0.005359507631510496
step: 110, loss: 0.04680558666586876
step: 120, loss: 0.017624638974666595
step: 130, loss: 0.008733206428587437
step: 140, loss: 0.0001783481566235423
step: 150, loss: 0.006963664665818214
step: 160, loss: 0.005737064406275749
step: 170, loss: 0.0015707053244113922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8608923884514436, f1=0.8839506172839506,
best_f1=0.8839506172839506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005544731975533068
step: 10, loss: 0.00032045828993432224
step: 20, loss: 0.0005893669440411031
step: 30, loss: 0.04557646065950394
step: 40, loss: 0.0032777825836092234
step: 50, loss: 0.023850150406360626
step: 60, loss: 0.0003636873443610966
step: 70, loss: 0.07192108035087585
step: 80, loss: 0.0018384740687906742
step: 90, loss: 0.0011244534980505705
step: 100, loss: 0.016579492017626762
step: 110, loss: 0.00045207151561044157
step: 120, loss: 0.007947679609060287
step: 130, loss: 0.0007796496502123773
step: 140, loss: 0.05421833693981171
step: 150, loss: 0.00026740357861854136
step: 160, loss: 0.0004414777795318514
step: 170, loss: 0.00018746277783066034
/usr/local/lib/python3.7/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=0.8557213930348259, f1=0.8738317757009346,
best_f1=0.8839506172839506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0043248082511126995
step: 10, loss: 0.008128384128212929

```

```

step: 20, loss: 0.010424666106700897
step: 30, loss: 0.0006804185104556382
step: 40, loss: 0.0019060930935665965
step: 50, loss: 0.01069556549191475
step: 60, loss: 0.0004408438690006733
step: 70, loss: 0.003092111088335514
step: 80, loss: 0.00031300695263780653
step: 90, loss: 0.0001090450314222835
step: 100, loss: 0.0002080088888760656
step: 110, loss: 0.024203820154070854
step: 120, loss: 0.0011012288741767406
step: 130, loss: 0.02689252607524395
step: 140, loss: 0.047863177955150604
step: 150, loss: 0.030882207676768303
step: 160, loss: 0.00011426792480051517
step: 170, loss: 0.02880263887345791
/usr/local/lib/python3.7/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=0.860759493670886, f1=0.8741092636579572,
best_f1=0.8839506172839506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.32126461318694e-05
step: 10, loss: 0.0001227132452186197
step: 20, loss: 0.00014808292326051742
step: 30, loss: 0.0004802976327482611
step: 40, loss: 0.00021764312987215817
step: 50, loss: 0.0005466660368256271
step: 60, loss: 0.00032181263668462634
step: 70, loss: 0.05546221882104874
step: 80, loss: 0.0009665468242019415
step: 90, loss: 0.0014557268004864454
step: 100, loss: 0.0632845014333725
step: 110, loss: 0.00010650113108567894
step: 120, loss: 0.00019878796592820436
step: 130, loss: 0.00033703012741170824
step: 140, loss: 0.00999370776116848
step: 150, loss: 0.00013495705206878483
step: 160, loss: 0.00014273250417318195
step: 170, loss: 0.00010700136044761166
/usr/local/lib/python3.7/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=0.8548812664907651, f1=0.8922305764411028,
best_f1=0.8839506172839506
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.")
```

1916it [00:09, 196.87it/s]

load_f1 = 0.5348101265822786

real_f1 = 0.49786628733997157

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

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

4398it [00:34, 127.44it/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: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5985886454582214
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.46306318044662476
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: 0.5330104827880859
step: 30, loss: 0.286662757396698
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 40, loss: 0.29351237416267395
step: 50, loss: 0.3039432466030121
step: 60, loss: 0.1480686217546463
step: 70, loss: 0.024733085185289383
step: 80, loss: 0.11752719432115555
step: 90, loss: 0.1181751936674118
step: 100, loss: 0.15450125932693481
step: 110, loss: 0.10614676773548126
step: 120, loss: 0.0428139790892601
step: 130, loss: 0.0049878372810781
step: 140, loss: 0.018554581329226494
step: 150, loss: 0.20293457806110382
step: 160, loss: 0.02337982878088951
step: 170, loss: 0.10706855356693268
step: 180, loss: 0.04962487518787384
step: 190, loss: 0.016780205070972443
step: 200, loss: 0.04288880527019501
step: 210, loss: 0.02750076726078987
step: 220, loss: 0.052833192050457
step: 230, loss: 0.002177477115765214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9149700598802396, f1=0.8978102189781021,
best_f1=0.8978102189781021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01627287268638611
step: 10, loss: 0.14999212324619293

```

```

step: 20, loss: 0.0271591953933239
step: 30, loss: 0.08628593385219574
step: 40, loss: 0.05806785076856613
step: 50, loss: 0.0588834211230278
step: 60, loss: 0.010238662362098694
step: 70, loss: 0.022965367883443832
step: 80, loss: 0.01564101129770279
step: 90, loss: 0.05784999579191208
step: 100, loss: 0.29518166184425354
step: 110, loss: 0.08612679690122604
step: 120, loss: 0.02806514874100685
step: 130, loss: 0.028243282809853554
step: 140, loss: 0.003990319557487965
step: 150, loss: 0.10721385478973389
step: 160, loss: 0.0253288671374321
step: 170, loss: 0.01961570978164673
step: 180, loss: 0.053713541477918625
step: 190, loss: 0.05498732626438141
step: 200, loss: 0.044068533927202225
step: 210, loss: 0.01995527371764183
step: 220, loss: 0.011401318944990635
step: 230, loss: 0.003799180733039975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9574944071588367, f1=0.962962962962963,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012411462143063545
step: 10, loss: 0.002607067348435521
step: 20, loss: 0.019294165074825287
step: 30, loss: 0.001347207697108388
step: 40, loss: 0.03425227850675583
step: 50, loss: 0.03075157105922699
step: 60, loss: 0.015392200089991093
step: 70, loss: 0.0023121926933526993
step: 80, loss: 0.00709372665733099
step: 90, loss: 0.013475550338625908
step: 100, loss: 0.0038747303187847137
step: 110, loss: 0.007597227115184069
step: 120, loss: 0.0023536267690360546
step: 130, loss: 0.019697029143571854

```

```

step: 140, loss: 0.0019056695746257901
step: 150, loss: 0.06866608560085297
step: 160, loss: 0.003727805335074663
step: 170, loss: 0.00113354017958045
step: 180, loss: 0.016036417335271835
step: 190, loss: 0.0872882679104805
step: 200, loss: 0.0171258132904768
step: 210, loss: 0.00137514213565737
step: 220, loss: 0.1560407280921936
step: 230, loss: 0.006181852426379919
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9657534246575342, f1=0.9643268124280783,
best_f1=0.9643268124280783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01969957910478115
step: 10, loss: 0.006231342442333698
step: 20, loss: 0.004555392544716597
step: 30, loss: 0.001496911165304482
step: 40, loss: 0.1415632665157318
step: 50, loss: 0.01932593807578087
step: 60, loss: 0.010552912019193172
step: 70, loss: 0.011098195798695087
step: 80, loss: 0.14663784205913544
step: 90, loss: 0.04890132322907448
step: 100, loss: 0.00965535081923008
step: 110, loss: 0.0017596465768292546
step: 120, loss: 0.005950122606009245
step: 130, loss: 0.0050640893168747425
step: 140, loss: 0.0009108344092965126
step: 150, loss: 0.004863319452852011
step: 160, loss: 0.001624571974389255
step: 170, loss: 0.0015576810110360384
step: 180, loss: 0.020280689001083374
step: 190, loss: 0.0010448622051626444
step: 200, loss: 0.04160595312714577
step: 210, loss: 0.0031488812528550625
step: 220, loss: 0.0013596814824268222
step: 230, loss: 0.0131192896515131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776286353467561, f1=0.9741863075196409,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02007640339434147
step: 10, loss: 0.06305708736181259
step: 20, loss: 0.01688235253095627
step: 30, loss: 0.003460424253717065
step: 40, loss: 0.0025789320934563875
step: 50, loss: 0.0018607917008921504
step: 60, loss: 0.013165310956537724
step: 70, loss: 0.004050331190228462
step: 80, loss: 0.033575426787137985
step: 90, loss: 0.09489700198173523
step: 100, loss: 0.00045996333938091993
step: 110, loss: 0.0016206272412091494
step: 120, loss: 0.0034181023947894573
step: 130, loss: 0.0008435574709437788
step: 140, loss: 0.005202909000217915
step: 150, loss: 0.029038207605481148
step: 160, loss: 0.0009167662938125432
step: 170, loss: 0.0059501961804926395
step: 180, loss: 0.006431016605347395
step: 190, loss: 0.009103849530220032
step: 200, loss: 0.031638868153095245
step: 210, loss: 0.024466656148433685
step: 220, loss: 0.005135828163474798
step: 230, loss: 0.014336689375340939
/usr/local/lib/python3.7/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=0.9755011135857461, f1=0.9700332963374029,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028446523938328028
step: 10, loss: 0.015812106430530548
step: 20, loss: 0.0068220579996705055
step: 30, loss: 0.001063308329321444
step: 40, loss: 0.0003127309901174158
step: 50, loss: 0.00043147685937583447

```

```

step: 60, loss: 0.0005128369084559381
step: 70, loss: 0.04843166097998619
step: 80, loss: 0.0009814901277422905
step: 90, loss: 0.0013747758930549026
step: 100, loss: 0.0004698718257714063
step: 110, loss: 0.006864911410957575
step: 120, loss: 0.0002856762148439884
step: 130, loss: 0.0030723202507942915
step: 140, loss: 0.0007603777339681983
step: 150, loss: 0.0021350590977817774
step: 160, loss: 0.004269353114068508
step: 170, loss: 0.0014365790411829948
step: 180, loss: 0.10076618939638138
step: 190, loss: 0.002474892418831587
step: 200, loss: 0.008323747664690018
step: 210, loss: 0.006384334992617369
step: 220, loss: 0.010522534139454365
step: 230, loss: 0.0061577060259878635
/usr/local/lib/python3.7/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=0.9619686800894856, f1=0.9731543624161074,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010730723151937127
step: 10, loss: 0.0006850440986454487
step: 20, loss: 0.027658309787511826
step: 30, loss: 0.00024993502302095294
step: 40, loss: 0.0003893952234648168
step: 50, loss: 0.003260352648794651
step: 60, loss: 0.005013057962059975
step: 70, loss: 0.000273840909358114
step: 80, loss: 0.0050417156890034676
step: 90, loss: 0.03383411839604378
step: 100, loss: 0.0019901618361473083
step: 110, loss: 0.0011121805291622877
step: 120, loss: 0.003421211149543524
step: 130, loss: 0.0003168702241964638
step: 140, loss: 0.0002553076483309269
step: 150, loss: 0.011100836098194122
step: 160, loss: 0.0029735290445387363
step: 170, loss: 0.00410822918638587
step: 180, loss: 0.000441730982856825
step: 190, loss: 0.0007428745157085359
step: 200, loss: 0.00658611673861742
step: 210, loss: 0.0017083996208384633

```

```

step: 220, loss: 0.00030407533631660044
step: 230, loss: 0.0005453271442092955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9796380090497738, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000815963780041784
step: 10, loss: 0.008249889127910137
step: 20, loss: 0.0015546535141766071
step: 30, loss: 0.0031125389505177736
step: 40, loss: 0.0012979070888832211
step: 50, loss: 0.0015546795912086964
step: 60, loss: 0.0020421885419636965
step: 70, loss: 0.00037753424840047956
step: 80, loss: 0.050137005746364594
step: 90, loss: 0.0006076554418541491
step: 100, loss: 0.0031380050349980593
step: 110, loss: 0.000796169217210263
step: 120, loss: 0.004445514176040888
step: 130, loss: 0.0005211663083173335
step: 140, loss: 0.0004829193640034646
step: 150, loss: 0.22163750231266022
step: 160, loss: 0.005198764614760876
step: 170, loss: 0.02795892208814621
step: 180, loss: 0.00017277240112889558
step: 190, loss: 0.0005425756680779159
step: 200, loss: 0.010969540104269981
step: 210, loss: 0.00038439620402641594
step: 220, loss: 0.0005056491354480386
step: 230, loss: 0.0002896322403103113
/usr/local/lib/python3.7/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=0.9784824462061155, f1=0.979591836734694,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028423259500414133
step: 10, loss: 0.0014163011219352484

```

```

step: 20, loss: 0.0002903113199863583
step: 30, loss: 0.00026004586834460497
step: 40, loss: 0.0014923388371244073
step: 50, loss: 0.00022458849707618356
step: 60, loss: 0.00022828762303106487
step: 70, loss: 0.045799750834703445
step: 80, loss: 0.00010816305439220741
step: 90, loss: 0.019359290599822998
step: 100, loss: 0.00043718202505260706
step: 110, loss: 0.0022449889220297337
step: 120, loss: 0.011164638213813305
step: 130, loss: 0.000772716011852026
step: 140, loss: 0.00016265457088593394
step: 150, loss: 0.00092192308511585
step: 160, loss: 0.0004268009215593338
step: 170, loss: 0.00012926143244840205
step: 180, loss: 0.0015216137981042266
step: 190, loss: 0.00012046322808600962
step: 200, loss: 0.00012861908180639148
step: 210, loss: 0.0006756072980351746
step: 220, loss: 0.00021993581322021782
step: 230, loss: 0.08754871785640717
/usr/local/lib/python3.7/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=0.9624724061810155, f1=0.9765886287625419,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016778863966464996
step: 10, loss: 0.00037420191802084446
step: 20, loss: 0.00048035054351203144
step: 30, loss: 0.00015489807992707938
step: 40, loss: 0.00023853952006902546
step: 50, loss: 0.00043863937025889754
step: 60, loss: 0.00014157046098262072
step: 70, loss: 0.03448071330785751
step: 80, loss: 0.018090903759002686
step: 90, loss: 0.00022885591897647828
step: 100, loss: 0.00014983126311562955
step: 110, loss: 0.012865434400737286
step: 120, loss: 0.00679759681224823
step: 130, loss: 0.0007455834420397878
step: 140, loss: 0.008563276380300522
step: 150, loss: 0.011028873734176159
step: 160, loss: 5.3682440920965746e-05
step: 170, loss: 0.00015051162336021662

```

```

step: 180, loss: 0.0031152626033872366
step: 190, loss: 0.00011920041288249195
step: 200, loss: 0.0003301123797427863
step: 210, loss: 0.02013220079243183
step: 220, loss: 0.0004362268082331866
step: 230, loss: 0.00010047078103525564
/usr/local/lib/python3.7/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=0.9788182831661093, f1=0.980963045912654,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027139464509673417
step: 10, loss: 0.00015273690223693848
step: 20, loss: 0.00016063866496551782
step: 30, loss: 0.00016327247431036085
step: 40, loss: 4.867949246545322e-05
step: 50, loss: 9.457142004976049e-05
step: 60, loss: 0.004468185361474752
step: 70, loss: 9.723375114845112e-05
step: 80, loss: 0.00019267664174549282
step: 90, loss: 0.2345869094133377
step: 100, loss: 0.00017037078214343637
step: 110, loss: 0.0013397865695878863
step: 120, loss: 0.0001656476379139349
step: 130, loss: 0.00012304879783187062
step: 140, loss: 0.005005016457289457
step: 150, loss: 0.00023483438417315483
step: 160, loss: 0.02034018561244011
step: 170, loss: 0.00025992622249759734
step: 180, loss: 0.00010282426228513941
step: 190, loss: 0.0001172442571260035
step: 200, loss: 0.0039676944725215435
step: 210, loss: 0.00022805901244282722
step: 220, loss: 0.000103449696325697
step: 230, loss: 0.00044158229138702154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.98, f1=0.977728285077951, best_f1=0.977728285077951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008410818409174681
step: 10, loss: 6.557380402227864e-05
step: 20, loss: 0.028813021257519722
step: 30, loss: 0.06865083426237106
step: 40, loss: 0.0001596102665644139
step: 50, loss: 0.0005379498470574617
step: 60, loss: 0.000154197245137766
step: 70, loss: 8.65873007569462e-05
step: 80, loss: 6.256745109567419e-05
step: 90, loss: 0.004183394834399223
step: 100, loss: 7.367042417172343e-05
step: 110, loss: 6.414650852093473e-05
step: 120, loss: 0.00016223193961195648
step: 130, loss: 9.104837954510003e-05
step: 140, loss: 0.001821825746446848
step: 150, loss: 0.00013479182962328196
step: 160, loss: 0.0001489101123297587
step: 170, loss: 0.00011202024325029925
step: 180, loss: 0.00014458954683505
step: 190, loss: 0.0001917419722303748
step: 200, loss: 0.00010403033229522407
step: 210, loss: 0.02727489359676838
step: 220, loss: 0.016499582678079605
step: 230, loss: 0.0006047657225281
/usr/local/lib/python3.7/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=0.9799554565701558, f1=0.9788182831661093,
best_f1=0.977728285077951
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001241095014847815
step: 10, loss: 0.00013642373960465193
step: 20, loss: 0.00011130441271234304
step: 30, loss: 0.004327305592596531
step: 40, loss: 0.00016309870989061892
step: 50, loss: 0.0003081504546571523
step: 60, loss: 0.0001675776147749275
step: 70, loss: 0.00036660776822827756
step: 80, loss: 7.482459477614611e-05
step: 90, loss: 8.711248665349558e-05
step: 100, loss: 0.00013556479825638235
step: 110, loss: 0.00013453661813400686
step: 120, loss: 0.0001050275459419936
step: 130, loss: 0.00014968551113270223
step: 140, loss: 5.013008922105655e-05

```

```

step: 150, loss: 0.002097709570080042
step: 160, loss: 0.03330523893237114
step: 170, loss: 0.00012375482765492052
step: 180, loss: 0.024253888055682182
step: 190, loss: 0.00012805026199202985
step: 200, loss: 2.6850646463572048e-05
step: 210, loss: 0.001983985072001815
step: 220, loss: 0.00015247594274114817
step: 230, loss: 0.0002748777042143047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9807909604519773, 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: 0.00011868815636262298
step: 10, loss: 5.5589491239516065e-05
step: 20, loss: 0.005226449575275183
step: 30, loss: 0.0006334580830298364
step: 40, loss: 6.094996570027433e-05
step: 50, loss: 6.156672316137701e-05
step: 60, loss: 7.699814159423113e-05
step: 70, loss: 7.52197956899181e-05
step: 80, loss: 8.375269680982456e-05
step: 90, loss: 0.0006279073422774673
step: 100, loss: 9.636614413466305e-05
step: 110, loss: 0.0001836169249145314
step: 120, loss: 6.385659071383998e-05
step: 130, loss: 0.0005648764781653881
step: 140, loss: 0.00014805154933128506
step: 150, loss: 0.0012301098322495818
step: 160, loss: 0.00014058996748644859
step: 170, loss: 0.00011061684199376032
step: 180, loss: 0.0001524124527350068
step: 190, loss: 6.582154310308397e-05
step: 200, loss: 0.0007181994733400643
step: 210, loss: 8.99505103006959e-05
step: 220, loss: 0.00010368631774326786
step: 230, loss: 0.0002478061942383647
/usr/local/lib/python3.7/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=0.9797752808988766, 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: 0.011604923754930496
step: 10, loss: 9.505688649369404e-05
step: 20, loss: 0.00025337835540995
step: 30, loss: 9.590038825990632e-05
step: 40, loss: 7.008015381870791e-05
step: 50, loss: 7.311425724765286e-05
step: 60, loss: 0.02904650755226612
step: 70, loss: 0.0001918633934110403
step: 80, loss: 0.0007835940341465175
step: 90, loss: 8.282549970317632e-05
step: 100, loss: 5.142491136211902e-05
step: 110, loss: 0.0001053876185324043
step: 120, loss: 0.07164055854082108
step: 130, loss: 7.769552757963538e-05
step: 140, loss: 0.0067395796068012714
step: 150, loss: 0.00031278160167858005
step: 160, loss: 0.023693829774856567
step: 170, loss: 5.093475556350313e-05
step: 180, loss: 0.00011295975127723068
step: 190, loss: 0.0002623202162794769
step: 200, loss: 0.0018368775490671396
step: 210, loss: 0.0076788500882685184
step: 220, loss: 5.985128154861741e-05
step: 230, loss: 7.377401925623417e-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=0.978865406006674, f1=0.9788182831661093,
best_f1=0.979591836734694
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, 138.42it/s]
load_f1 = 0.9810055865921787
real_f1 = 0.9799554565701558
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:34, 128.44it/s]

```

3.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

```
# Step 3: run the matcher
```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6513508558273315
```

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

```
step: 10, loss: 0.42886075377464294
```

```
step: 20, loss: 0.3785828948020935
```

```
step: 30, loss: 0.33199360966682434
```

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

```
step: 40, loss: 0.39993423223495483
```

```
step: 50, loss: 0.6168199181556702
```

```
step: 60, loss: 0.2533594071865082
```

step: 70, loss: 0.3169642388820648
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
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 80, loss: 0.6531227827072144
step: 90, loss: 0.14308220148086548
step: 100, loss: 0.1950315237045288
step: 110, loss: 0.13637728989124298
step: 120, loss: 0.13478681445121765
step: 130, loss: 0.09778774529695511
step: 140, loss: 0.25003325939178467
step: 150, loss: 0.04125254601240158
step: 160, loss: 0.11637409776449203
step: 170, loss: 0.16772107779979706
step: 180, loss: 0.039553090929985046
step: 190, loss: 0.06548202782869339
step: 200, loss: 0.07512587308883667
step: 210, loss: 0.02655310370028019
step: 220, loss: 0.16443580389022827
step: 230, loss: 0.14838692545890808
step: 240, loss: 0.017397165298461914
step: 250, loss: 0.09439046680927277
step: 260, loss: 0.08769006282091141
step: 270, loss: 0.3371157944202423
step: 280, loss: 0.04388100281357765
step: 290, loss: 0.23169025778770447
step: 300, loss: 0.04458032548427582
step: 310, loss: 0.1368204951286316
step: 320, loss: 0.03079783171415329
step: 330, loss: 0.11395791172981262
step: 340, loss: 0.3693923056125641
step: 350, loss: 0.10560955852270126
step: 360, loss: 0.08441122621297836
step: 370, loss: 0.08038989454507828
step: 380, loss: 0.13195286691188812
step: 390, loss: 0.03598331660032272
step: 400, loss: 0.07602386921644211
step: 410, loss: 0.26666516065597534
step: 420, loss: 0.055200766772031784
step: 430, loss: 0.006116804666817188
step: 440, loss: 0.05614319071173668
step: 450, loss: 0.01695256493985653
step: 460, loss: 0.00760037312284112
step: 470, loss: 0.054412711411714554
step: 480, loss: 0.09811929613351822
step: 490, loss: 0.1445510983467102
step: 500, loss: 0.09490140527486801
step: 510, loss: 0.04693603515625

```

step: 520, loss: 0.05753111094236374
step: 530, loss: 0.04665924608707428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9428044280442804, f1=0.9459584295612009,
best_f1=0.9459584295612009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05078479275107384
step: 10, loss: 0.08059235662221909
step: 20, loss: 0.02388208545744419
step: 30, loss: 0.05734280124306679
step: 40, loss: 0.0809810608625412
step: 50, loss: 0.05428146943449974
step: 60, loss: 0.1372581422328949
step: 70, loss: 0.02618948370218277
step: 80, loss: 0.024614308029413223
step: 90, loss: 0.00577305955812335
step: 100, loss: 0.14990083873271942
step: 110, loss: 0.024329250678420067
step: 120, loss: 0.2571744918823242
step: 130, loss: 0.01863725483417511
step: 140, loss: 0.016882121562957764
step: 150, loss: 0.025902200490236282
step: 160, loss: 0.03839440271258354
step: 170, loss: 0.01604059338569641
step: 180, loss: 0.006381090264767408
step: 190, loss: 0.018150871619582176
step: 200, loss: 0.26779428124427795
step: 210, loss: 0.015470638871192932
step: 220, loss: 0.0018898320849984884
step: 230, loss: 0.030588999390602112
step: 240, loss: 0.05376714840531349
step: 250, loss: 0.021073147654533386
step: 260, loss: 0.04236188158392906
step: 270, loss: 0.00656430097296834
step: 280, loss: 0.05755668878555298
step: 290, loss: 0.026543308049440384
step: 300, loss: 0.06997933983802795
step: 310, loss: 0.06925717741250992
step: 320, loss: 0.0536635553269386
step: 330, loss: 0.019069254398345947

```

```

step: 340, loss: 0.04224452003836632
step: 350, loss: 0.007727998308837414
step: 360, loss: 0.08531522750854492
step: 370, loss: 0.002093919552862644
step: 380, loss: 0.1064850389957428
step: 390, loss: 0.022435838356614113
step: 400, loss: 0.11934297531843185
step: 410, loss: 0.01065819337964058
step: 420, loss: 0.053320739418268204
step: 430, loss: 0.23717407882213593
step: 440, loss: 0.023056287318468094
step: 450, loss: 0.020341821014881134
step: 460, loss: 0.07129845768213272
step: 470, loss: 0.006932047661393881
step: 480, loss: 0.005479946732521057
step: 490, loss: 0.02566635236144066
step: 500, loss: 0.004823415540158749
step: 510, loss: 0.01107735000550747
step: 520, loss: 0.35933366417884827
step: 530, loss: 0.08912584185600281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9478976234003655, f1=0.9441391941391942,
best_f1=0.9441391941391942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05614538490772247
step: 10, loss: 0.021772734820842743
step: 20, loss: 0.011381574906408787
step: 30, loss: 0.029036615043878555
step: 40, loss: 0.014506028033792973
step: 50, loss: 0.008777650073170662
step: 60, loss: 0.0239475816488266
step: 70, loss: 0.00803028792142868
step: 80, loss: 0.04620478302240372
step: 90, loss: 0.01429987046867609
step: 100, loss: 0.12553662061691284
step: 110, loss: 0.0525994598865509
step: 120, loss: 0.02684818021953106
step: 130, loss: 0.040099307894706726
step: 140, loss: 0.007729605305939913
step: 150, loss: 0.011131574399769306

```

```

step: 160, loss: 0.03837612643837929
step: 170, loss: 0.015010060742497444
step: 180, loss: 0.0062378342263400555
step: 190, loss: 0.013164043426513672
step: 200, loss: 0.11807551234960556
step: 210, loss: 0.0343744121491909
step: 220, loss: 0.12866033613681793
step: 230, loss: 0.02630254253745079
step: 240, loss: 0.04492072016000748
step: 250, loss: 0.056344084441661835
step: 260, loss: 0.12242460995912552
step: 270, loss: 0.031976863741874695
step: 280, loss: 0.017148789018392563
step: 290, loss: 0.00477214390411973
step: 300, loss: 0.11134250462055206
step: 310, loss: 0.04703102260828018
step: 320, loss: 0.11182492971420288
step: 330, loss: 0.02134961262345314
step: 340, loss: 0.007049191277474165
step: 350, loss: 0.132350891828537
step: 360, loss: 0.026074176654219627
step: 370, loss: 0.04670252278447151
step: 380, loss: 0.0016285099554806948
step: 390, loss: 0.0027198500465601683
step: 400, loss: 0.16355179250240326
step: 410, loss: 0.0781225934624672
step: 420, loss: 0.022187262773513794
step: 430, loss: 0.06688586622476578
step: 440, loss: 0.196445494890213
step: 450, loss: 0.0772760659456253
step: 460, loss: 0.0313529372215271
step: 470, loss: 0.06376072019338608
step: 480, loss: 0.08668465167284012
step: 490, loss: 0.04767686873674393
step: 500, loss: 0.005084974691271782
step: 510, loss: 0.04389417916536331
step: 520, loss: 0.002752171363681555
step: 530, loss: 0.00872766226530075
/usr/local/lib/python3.7/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=0.947565543071161, f1=0.9394081728511038,
best_f1=0.9441391941391942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02354799211025238
step: 10, loss: 0.01132939476519823

```

step: 20, loss: 0.1446138173341751
step: 30, loss: 0.12955638766288757
step: 40, loss: 0.007413613144308329
step: 50, loss: 0.018145272508263588
step: 60, loss: 0.002228720346465707
step: 70, loss: 0.04309928044676781
step: 80, loss: 0.0029626006726175547
step: 90, loss: 0.15200214087963104
step: 100, loss: 0.008765032514929771
step: 110, loss: 0.14883890748023987
step: 120, loss: 0.00209300359711051
step: 130, loss: 0.01598373055458069
step: 140, loss: 0.0725807324051857
step: 150, loss: 0.03272542357444763
step: 160, loss: 0.046492889523506165
step: 170, loss: 0.05231134593486786
step: 180, loss: 0.11360405385494232
step: 190, loss: 0.04338717833161354
step: 200, loss: 0.08136431872844696
step: 210, loss: 0.0003238725184928626
step: 220, loss: 0.0032613978255540133
step: 230, loss: 0.004486010875552893
step: 240, loss: 0.051580753177404404
step: 250, loss: 0.10055742412805557
step: 260, loss: 0.0022741584107279778
step: 270, loss: 0.028877301141619682
step: 280, loss: 0.017032526433467865
step: 290, loss: 0.07918445020914078
step: 300, loss: 0.0009013751405291259
step: 310, loss: 0.005216993857175112
step: 320, loss: 0.06909215450286865
step: 330, loss: 0.08327000588178635
step: 340, loss: 0.008655417710542679
step: 350, loss: 0.07685662060976028
step: 360, loss: 0.03483111783862114
step: 370, loss: 0.004678904078900814
step: 380, loss: 0.005445886868983507
step: 390, loss: 0.0002648108347784728
step: 400, loss: 0.026526160538196564
step: 410, loss: 0.002133092610165477
step: 420, loss: 0.006524431984871626
step: 430, loss: 0.0025103408843278885
step: 440, loss: 0.0007688040495850146
step: 450, loss: 0.00547578651458025
step: 460, loss: 0.013283107429742813
step: 470, loss: 0.001626454759389162
step: 480, loss: 0.010723773390054703
step: 490, loss: 0.0029622907750308514

```

step: 500, loss: 0.035415127873420715
step: 510, loss: 0.017172854393720627
step: 520, loss: 0.017566632479429245
step: 530, loss: 0.16223081946372986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9504587155963302, f1=0.9494668521094112,
best_f1=0.9494668521094112
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035888575948774815
step: 10, loss: 0.027175992727279663
step: 20, loss: 0.005421656649559736
step: 30, loss: 0.010305817238986492
step: 40, loss: 0.0049199857749044895
step: 50, loss: 0.15992295742034912
step: 60, loss: 0.044017113745212555
step: 70, loss: 0.004693233873695135
step: 80, loss: 0.008405561558902264
step: 90, loss: 0.07342749834060669
step: 100, loss: 0.12217896431684494
step: 110, loss: 0.015391725115478039
step: 120, loss: 0.18679948151111603
step: 130, loss: 0.011955457739531994
step: 140, loss: 0.00127657619304955
step: 150, loss: 0.0038639565464109182
step: 160, loss: 0.003645320888608694
step: 170, loss: 0.10611387342214584
step: 180, loss: 0.012584797106683254
step: 190, loss: 0.005061610136181116
step: 200, loss: 0.005655793007463217
step: 210, loss: 0.005057953298091888
step: 220, loss: 0.005245058797299862
step: 230, loss: 0.0016978151397779584
step: 240, loss: 0.004919727798551321
step: 250, loss: 0.12561218440532684
step: 260, loss: 0.0012892138911411166
step: 270, loss: 0.008875640109181404
step: 280, loss: 0.005590060260146856
step: 290, loss: 0.008068367838859558
step: 300, loss: 0.16513396799564362
step: 310, loss: 0.06729307025671005

```



```

step: 320, loss: 0.03687531128525734
step: 330, loss: 0.0027777724899351597
step: 340, loss: 0.04536738619208336
step: 350, loss: 0.0020580908749252558
step: 360, loss: 0.0012888186611235142
step: 370, loss: 0.002730901585891843
step: 380, loss: 0.0021176431328058243
step: 390, loss: 0.01913992315530777
step: 400, loss: 0.005788277368992567
step: 410, loss: 0.06400617212057114
step: 420, loss: 0.22724109888076782
step: 430, loss: 0.0508260652422905
step: 440, loss: 0.0009911073138937354
step: 450, loss: 0.01769956946372986
step: 460, loss: 0.026599029079079628
step: 470, loss: 0.05785210803151131
step: 480, loss: 0.016791783273220062
step: 490, loss: 0.0025024297647178173
step: 500, loss: 0.020099051296710968
step: 510, loss: 0.0038694622926414013
step: 520, loss: 0.08150061964988708
step: 530, loss: 0.035810042172670364
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9537037037037037, f1=0.9465861588481189,
best_f1=0.9465861588481189
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011633873917162418
step: 10, loss: 0.00038046675035730004
step: 20, loss: 0.02810082584619522
step: 30, loss: 0.0005171233788132668
step: 40, loss: 0.0002543991431593895
step: 50, loss: 0.00019060110207647085
step: 60, loss: 0.004893911071121693
step: 70, loss: 0.0005140131106600165
step: 80, loss: 0.00022765759786125273
step: 90, loss: 0.0005395900225266814
step: 100, loss: 0.016393527388572693
step: 110, loss: 0.010245993733406067
step: 120, loss: 0.023041611537337303
step: 130, loss: 0.01024584099650383

```

```

step: 140, loss: 0.005393208935856819
step: 150, loss: 0.00031360238790512085
step: 160, loss: 0.17022724449634552
step: 170, loss: 0.011240157298743725
step: 180, loss: 0.0011718993773683906
step: 190, loss: 0.05520764738321304
step: 200, loss: 0.003734109690412879
step: 210, loss: 0.0018803102429956198
step: 220, loss: 0.005117360502481461
step: 230, loss: 0.005418921355158091
step: 240, loss: 0.00031943846261128783
step: 250, loss: 0.026300178840756416
step: 260, loss: 0.00020136513921897858
step: 270, loss: 0.0001925229444168508
step: 280, loss: 0.0379154346883297
step: 290, loss: 0.0012607863172888756
step: 300, loss: 0.002601556945592165
step: 310, loss: 0.019063735380768776
step: 320, loss: 0.0006042185123078525
step: 330, loss: 0.09199431538581848
step: 340, loss: 0.000643911596853286
step: 350, loss: 0.012028411030769348
step: 360, loss: 0.03442425653338432
step: 370, loss: 0.006517614237964153
step: 380, loss: 0.0018213248113170266
step: 390, loss: 0.0009769309544935822
step: 400, loss: 0.012520739808678627
step: 410, loss: 0.0011948631145060062
step: 420, loss: 0.015055187046527863
step: 430, loss: 0.0034673188347369432
step: 440, loss: 0.0035787078086286783
step: 450, loss: 0.12685106694698334
step: 460, loss: 0.005751608870923519
step: 470, loss: 0.0024628164246678352
step: 480, loss: 0.006628271192312241
step: 490, loss: 0.01780719682574272
step: 500, loss: 0.007182836998254061
step: 510, loss: 0.20663690567016602
step: 520, loss: 0.0013723638840019703
step: 530, loss: 0.002622483763843775
/usr/local/lib/python3.7/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=0.9493908153701968, f1=0.9389493610979649,
best_f1=0.9465861588481189
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.0036254555452615023
step: 10, loss: 0.0016188056906685233
step: 20, loss: 0.0011818923521786928
step: 30, loss: 0.008901229128241539
step: 40, loss: 0.009566846303641796
step: 50, loss: 0.0019430492538958788
step: 60, loss: 0.0010474780574440956
step: 70, loss: 0.04974432662129402
step: 80, loss: 0.013374142348766327
step: 90, loss: 6.931550160516053e-05
step: 100, loss: 0.09121038019657135
step: 110, loss: 0.00014568821643479168
step: 120, loss: 0.0017947289161384106
step: 130, loss: 0.0002127937914337963
step: 140, loss: 0.00024666846729815006
step: 150, loss: 0.00019373142276890576
step: 160, loss: 0.0014037755317986012
step: 170, loss: 0.00283612753264606
step: 180, loss: 0.0027672548312693834
step: 190, loss: 0.0017725098878145218
step: 200, loss: 0.0003419609274715185
step: 210, loss: 0.019799938425421715
step: 220, loss: 0.00031877413857728243
step: 230, loss: 0.0051512098871171474
step: 240, loss: 0.01328946091234684
step: 250, loss: 0.0011884679552167654
step: 260, loss: 0.004089240450412035
step: 270, loss: 0.003382837399840355
step: 280, loss: 0.020110785961151123
step: 290, loss: 0.003770227311179042
step: 300, loss: 0.0008189742802642286
step: 310, loss: 0.0028167576529085636
step: 320, loss: 0.0012941996101289988
step: 330, loss: 0.0006240601069293916
step: 340, loss: 0.0038316049613058567
step: 350, loss: 0.010864685289561749
step: 360, loss: 0.0014703420456498861
step: 370, loss: 0.029963267967104912
step: 380, loss: 0.0072457194328308105
step: 390, loss: 0.002710535656660795
step: 400, loss: 0.008549729362130165
step: 410, loss: 0.000288380280835554
step: 420, loss: 0.061632465571165085
step: 430, loss: 0.009246785193681717
step: 440, loss: 0.0005015728529542685
step: 450, loss: 0.010660281404852867
step: 460, loss: 0.008173348382115364
step: 470, loss: 0.11304935067892075

```

step: 480, loss: 0.0033792161848396063
step: 490, loss: 0.02915484830737114
step: 500, loss: 0.0014358971966430545
step: 510, loss: 0.0010164218256250024
step: 520, loss: 0.0009084272314794362
step: 530, loss: 0.0036770938895642757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.954460966542751, f1=0.941340782122905,
best_f1=0.941340782122905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010032580466941
step: 10, loss: 0.0019038631580770016
step: 20, loss: 0.0018864102894440293
step: 30, loss: 0.0006176215247251093
step: 40, loss: 0.0004620479012373835
step: 50, loss: 0.0062746102921664715
step: 60, loss: 0.06634693592786789
step: 70, loss: 0.0008565267198719084
step: 80, loss: 0.003131016856059432
step: 90, loss: 0.0004976115887984633
step: 100, loss: 0.0026596239767968655
step: 110, loss: 0.000530813995283097
step: 120, loss: 0.0011359774507582188
step: 130, loss: 0.017739856615662575
step: 140, loss: 8.21769644971937e-05
step: 150, loss: 0.0015320780221372843
step: 160, loss: 0.00016825657803565264
step: 170, loss: 0.10173790156841278
step: 180, loss: 0.0013367165811359882
step: 190, loss: 0.013768619857728481
step: 200, loss: 0.004557306412607431
step: 210, loss: 0.03990795463323593
step: 220, loss: 0.0017330129630863667
step: 230, loss: 0.04092489555478096
step: 240, loss: 0.000538055261131376
step: 250, loss: 0.0005739468615502119
step: 260, loss: 0.00019444458303041756
step: 270, loss: 0.0077210976742208
step: 280, loss: 0.0007479431806132197
step: 290, loss: 0.00017730917898006737

```

```

step: 300, loss: 0.0007712225778959692
step: 310, loss: 0.0006063068867661059
step: 320, loss: 0.0015874541131779552
step: 330, loss: 0.0018247654661536217
step: 340, loss: 0.005642653908580542
step: 350, loss: 0.004564434289932251
step: 360, loss: 0.009012256748974323
step: 370, loss: 0.0746477022767067
step: 380, loss: 0.00038250256329774857
step: 390, loss: 0.0027754423208534718
step: 400, loss: 0.040914926677942276
step: 410, loss: 0.0026866637635976076
step: 420, loss: 0.0121089993044734
step: 430, loss: 0.02423671819269657
step: 440, loss: 0.0013615645002573729
step: 450, loss: 0.00048460846301168203
step: 460, loss: 0.0025868876837193966
step: 470, loss: 0.007606182247400284
step: 480, loss: 0.0018179566832259297
step: 490, loss: 0.0036869822070002556
step: 500, loss: 0.009920899756252766
step: 510, loss: 0.0011168535565957427
step: 520, loss: 0.0005379333742894232
step: 530, loss: 0.0059578861109912395
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9546296296296296, f1=0.9511854951185496,
best_f1=0.9511854951185496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005334251327440143
step: 10, loss: 0.002247432479634881
step: 20, loss: 0.0007585345301777124
step: 30, loss: 0.07441054284572601
step: 40, loss: 0.0003121541813015938
step: 50, loss: 0.001371685997582972
step: 60, loss: 0.001154415775090456
step: 70, loss: 0.1406290978193283
step: 80, loss: 0.013083349913358688
step: 90, loss: 0.009940076619386673
step: 100, loss: 0.012846402823925018
step: 110, loss: 0.03084787353873253

```

```
step: 120, loss: 0.002796806627884507
step: 130, loss: 0.004306696355342865
step: 140, loss: 0.05495356395840645
step: 150, loss: 0.00677513936534524
step: 160, loss: 0.004413714166730642
step: 170, loss: 0.0033152340911328793
step: 180, loss: 0.0038489929866045713
step: 190, loss: 0.003770094132050872
step: 200, loss: 0.00029866534168832004
step: 210, loss: 0.05463731288909912
step: 220, loss: 0.0008622459718026221
step: 230, loss: 0.021390054374933243
step: 240, loss: 0.001203826512210071
step: 250, loss: 0.001061864197254181
step: 260, loss: 0.001422010362148285
step: 270, loss: 0.06933692842721939
step: 280, loss: 0.018340038135647774
step: 290, loss: 0.0005460610846057534
step: 300, loss: 0.004144018981605768
step: 310, loss: 0.07821033895015717
step: 320, loss: 0.0019523435039445758
step: 330, loss: 0.005686059128493071
step: 340, loss: 0.00174853450153023
step: 350, loss: 0.004181536380201578
step: 360, loss: 0.00014796122559346259
step: 370, loss: 0.004157180432230234
step: 380, loss: 0.0020566813182085752
step: 390, loss: 0.0006421500002034009
step: 400, loss: 0.08076459914445877
step: 410, loss: 0.004288447089493275
step: 420, loss: 0.0002730978885665536
step: 430, loss: 0.008540279231965542
step: 440, loss: 0.0002345692046219483
step: 450, loss: 0.0010500096250325441
step: 460, loss: 0.00014782921061851084
step: 470, loss: 0.0005553120863623917
step: 480, loss: 0.0006084935157559812
step: 490, loss: 0.012781291268765926
step: 500, loss: 0.0004388696688693017
step: 510, loss: 7.807542715454474e-05
step: 520, loss: 0.00139120954554528
step: 530, loss: 0.0004020150809083134
/usr/local/lib/python3.7/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=0.9512761020881672, f1=0.946236559139785,
best_f1=0.9511854951185496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.005764072295278311
step: 10, loss: 0.0007877305615693331
step: 20, loss: 0.00021720287622883916
step: 30, loss: 0.0005981717840768397
step: 40, loss: 0.00018462465959601104
step: 50, loss: 0.000951346184592694
step: 60, loss: 0.0018953434191644192
step: 70, loss: 0.00016007143130991608
step: 80, loss: 0.0079400185495615
step: 90, loss: 0.0002991695364471525
step: 100, loss: 0.0031278470996767282
step: 110, loss: 0.0024565032217651606
step: 120, loss: 0.0009272571187466383
step: 130, loss: 0.0002166744670830667
step: 140, loss: 0.0007409299141727388
step: 150, loss: 0.0003114535356871784
step: 160, loss: 0.037262335419654846
step: 170, loss: 0.0003343233256600797
step: 180, loss: 0.00034866732312366366
step: 190, loss: 0.00013092035078443587
step: 200, loss: 0.00017869028670247644
step: 210, loss: 0.012779982760548592
step: 220, loss: 0.0004108510911464691
step: 230, loss: 0.0007175069767981768
step: 240, loss: 0.00016929973207879812
step: 250, loss: 0.00030057993717491627
step: 260, loss: 0.00020162966393399984
step: 270, loss: 0.00022287183674052358
step: 280, loss: 0.009275112301111221
step: 290, loss: 0.0003664171672426164
step: 300, loss: 0.03298614174127579
step: 310, loss: 0.026387309655547142
step: 320, loss: 0.010067138820886612
step: 330, loss: 0.0019904873333871365
step: 340, loss: 0.00042808891157619655
step: 350, loss: 0.0003807591856457293
step: 360, loss: 0.00012523798795882612
step: 370, loss: 0.007632396649569273
step: 380, loss: 0.002776411361992359
step: 390, loss: 0.0009426468750461936
step: 400, loss: 0.0019333807285875082
step: 410, loss: 0.0012728554429486394
step: 420, loss: 6.80371158523485e-05
step: 430, loss: 0.00022727457690052688
step: 440, loss: 3.9768226997694e-05
step: 450, loss: 0.00045489644980989397
```

```

step: 460, loss: 0.0007317003328353167
step: 470, loss: 0.005099378060549498
step: 480, loss: 0.0006706186686642468
step: 490, loss: 0.03442710265517235
step: 500, loss: 0.014191499911248684
step: 510, loss: 0.0001254235248779878
step: 520, loss: 0.008957354351878166
step: 530, loss: 0.0075304643251001835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9565627950897073, f1=0.9457659372026642,
best_f1=0.9457659372026642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007041034987196326
step: 10, loss: 0.0006722526159137487
step: 20, loss: 0.0018182715866714716
step: 30, loss: 0.00014218362048268318
step: 40, loss: 0.0001629685575608164
step: 50, loss: 0.00034360497375018895
step: 60, loss: 0.00013086361286696047
step: 70, loss: 0.00017551759083289653
step: 80, loss: 0.00017124437727034092
step: 90, loss: 0.010655885562300682
step: 100, loss: 0.00028909838874824345
step: 110, loss: 0.000348977540852502
step: 120, loss: 0.0007766594644635916
step: 130, loss: 0.000993122230283916
step: 140, loss: 0.00017585967725608498
step: 150, loss: 0.00042270065750926733
step: 160, loss: 0.0010242968564853072
step: 170, loss: 0.012350370176136494
step: 180, loss: 0.0005243096384219825
step: 190, loss: 0.00042356198537163436
step: 200, loss: 0.00012952891120221466
step: 210, loss: 0.0004654170770663768
step: 220, loss: 0.013884457759559155
step: 230, loss: 0.0003303339472040534
step: 240, loss: 0.002298501553013921
step: 250, loss: 0.0006729950546287
step: 260, loss: 0.003019292140379548
step: 270, loss: 0.0035666669718921185

```



```

step: 280, loss: 0.00014349511184263974
step: 290, loss: 0.00017553611542098224
step: 300, loss: 0.00020900575327686965
step: 310, loss: 0.000279203406535089
step: 320, loss: 0.0013365625636652112
step: 330, loss: 5.52647361473646e-05
step: 340, loss: 0.0010374158155173063
step: 350, loss: 0.00010350308730266988
step: 360, loss: 0.00025496663874946535
step: 370, loss: 0.0016540128272026777
step: 380, loss: 0.0005710582481697202
step: 390, loss: 0.0010808255756273866
step: 400, loss: 0.00035469213617034256
step: 410, loss: 0.0009547362569719553
step: 420, loss: 0.0031391591764986515
step: 430, loss: 0.0009751870529726148
step: 440, loss: 7.640311378054321e-05
step: 450, loss: 0.041168007999658585
step: 460, loss: 0.0003363917348906398
step: 470, loss: 0.0012885734904557467
step: 480, loss: 0.0004172933695372194
step: 490, loss: 3.4388420317554846e-05
step: 500, loss: 0.0032209299970418215
step: 510, loss: 0.00012936853454448283
step: 520, loss: 0.00022221433755476028
step: 530, loss: 0.012787282466888428
/usr/local/lib/python3.7/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=0.9555035128805621, f1=0.9441052137153593,
best_f1=0.9457659372026642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003276368952356279
step: 10, loss: 5.860593228135258e-05
step: 20, loss: 8.619054278824478e-05
step: 30, loss: 0.0001129853626480326
step: 40, loss: 0.0009533805423416197
step: 50, loss: 0.002669096924364567
step: 60, loss: 0.00042852701153606176
step: 70, loss: 0.0003016793925780803
step: 80, loss: 0.0008949902257882059
step: 90, loss: 0.00019386704661883414
step: 100, loss: 0.09040279686450958
step: 110, loss: 0.0002275906881550327
step: 120, loss: 0.0018803535494953394
step: 130, loss: 0.0013305670581758022

```

```

step: 140, loss: 0.002382010919973254
step: 150, loss: 0.001419675536453724
step: 160, loss: 0.0001925164251588285
step: 170, loss: 0.00022446928778663278
step: 180, loss: 4.8193251132033765e-05
step: 190, loss: 0.00010182059486396611
step: 200, loss: 0.0009592513088136911
step: 210, loss: 0.000546683615539223
step: 220, loss: 3.7879097362747416e-05
step: 230, loss: 2.7132522518513724e-05
step: 240, loss: 0.0005267211818136275
step: 250, loss: 4.723635720438324e-05
step: 260, loss: 0.00011862632527481765
step: 270, loss: 0.0017427988350391388
step: 280, loss: 0.011480770073831081
step: 290, loss: 0.0003190996067132801
step: 300, loss: 0.0010588295990601182
step: 310, loss: 4.11720247939229e-05
step: 320, loss: 0.004764589946717024
step: 330, loss: 0.0012221739161759615
step: 340, loss: 3.296904105809517e-05
step: 350, loss: 0.00018576171714812517
step: 360, loss: 6.887748168082908e-05
step: 370, loss: 1.3764516552328132e-05
step: 380, loss: 0.00019029121904168278
step: 390, loss: 6.693087198073044e-05
step: 400, loss: 2.137819137715269e-05
step: 410, loss: 0.010451452806591988
step: 420, loss: 5.7833454775391147e-05
step: 430, loss: 2.0916708308504894e-05
step: 440, loss: 0.00020297043374739587
step: 450, loss: 0.0009063658071681857
step: 460, loss: 8.212996908696368e-05
step: 470, loss: 0.001295882393606007
step: 480, loss: 0.0001301008596783504
step: 490, loss: 3.1030242098495364e-05
step: 500, loss: 8.662600885145366e-05
step: 510, loss: 0.0010075564496219158
step: 520, loss: 0.002689779968932271
step: 530, loss: 0.0002781909133773297
/usr/local/lib/python3.7/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=0.9511192325262678, f1=0.9460694698354661,
best_f1=0.9457659372026642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.017533449456095695
step: 10, loss: 0.0024513513781130314
step: 20, loss: 0.00010880371701205149
step: 30, loss: 1.9992843590443954e-05
step: 40, loss: 9.875890827970579e-05
step: 50, loss: 0.005862200167030096
step: 60, loss: 2.4972772735054605e-05
step: 70, loss: 0.00011304745567031205
step: 80, loss: 4.1601037082728e-05
step: 90, loss: 0.000587356451433152
step: 100, loss: 5.3155112254898995e-05
step: 110, loss: 1.599593451828696e-05
step: 120, loss: 2.734189183684066e-05
step: 130, loss: 0.00020842664525844157
step: 140, loss: 0.0028762714937329292
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.0006240860093384981
step: 160, loss: 0.001281054806895554
step: 170, loss: 0.006341150030493736
step: 180, loss: 4.016193997813389e-05
step: 190, loss: 0.0006620088825002313
step: 200, loss: 5.742501889471896e-05
step: 210, loss: 0.0012274577748030424
step: 220, loss: 0.02222985401749611
step: 230, loss: 0.0015245892573148012
step: 240, loss: 0.001800760393962264
step: 250, loss: 0.0002506155869923532
step: 260, loss: 0.00010655339428922161
step: 270, loss: 0.0001479134225519374
step: 280, loss: 0.00016709588817320764
step: 290, loss: 1.7884131011669524e-05
step: 300, loss: 6.0800055507570505e-05
step: 310, loss: 8.928225724957883e-05
step: 320, loss: 7.735688268439844e-05
step: 330, loss: 0.00013802939793094993
step: 340, loss: 0.0023106818553060293
step: 350, loss: 3.234377800254151e-05
step: 360, loss: 0.07218477129936218
step: 370, loss: 2.4084874894469976e-05
step: 380, loss: 4.0478145820088685e-05
step: 390, loss: 5.403832619776949e-05
step: 400, loss: 0.002286540111526847
step: 410, loss: 5.0178769015474245e-05
step: 420, loss: 3.468776776571758e-05
step: 430, loss: 8.024711132748052e-05
step: 440, loss: 0.0019442803459241986
step: 450, loss: 4.623828863259405e-05
step: 460, loss: 0.00018470916256774217

```

step: 470, loss: 0.0018612454878166318
step: 480, loss: 1.2170161426183768e-05
step: 490, loss: 4.854653525399044e-05
step: 500, loss: 0.0008894734783098102
step: 510, loss: 0.0005042511038482189
step: 520, loss: 0.00030825065914541483
step: 530, loss: 9.055546979652718e-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=0.9547363509099394, f1=0.9439252336448598,
best_f1=0.9457659372026642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004732597153633833
step: 10, loss: 5.5282820540014654e-05
step: 20, loss: 0.00011132177314721048
step: 30, loss: 0.0001229861081810668
step: 40, loss: 0.0005405173869803548
step: 50, loss: 7.015017763478681e-05
step: 60, loss: 3.692158861667849e-05
step: 70, loss: 4.2109521018574014e-05
step: 80, loss: 6.789184408262372e-05
step: 90, loss: 1.766098284861073e-05
step: 100, loss: 2.3125556253944524e-05
step: 110, loss: 4.443003854248673e-05
step: 120, loss: 2.159109862986952e-05
step: 130, loss: 2.1144074707990512e-05
step: 140, loss: 3.0672519642394036e-05
step: 150, loss: 0.00035689075593836606
step: 160, loss: 5.041106487624347e-05
step: 170, loss: 6.223942909855396e-05
step: 180, loss: 4.4937671191291884e-05
step: 190, loss: 0.013139468617737293
step: 200, loss: 6.723454862367362e-05
step: 210, loss: 0.00011698609159793705
step: 220, loss: 2.4215907615143806e-05
step: 230, loss: 1.8130387616110966e-05
step: 240, loss: 0.000719869218301028
step: 250, loss: 0.0019252797355875373
step: 260, loss: 2.9636325052706525e-05
step: 270, loss: 8.534274093108252e-05
step: 280, loss: 4.234533116687089e-05
step: 290, loss: 0.0015082692261785269
step: 300, loss: 0.0001305342884734273
step: 310, loss: 0.008891143836081028
step: 320, loss: 3.430727883824147e-05

```

```

step: 330, loss: 0.0011567905312404037
step: 340, loss: 0.0016592546598985791
step: 350, loss: 0.000385488587198779
step: 360, loss: 5.149958815309219e-05
step: 370, loss: 0.0004464008379727602
step: 380, loss: 0.009779959917068481
step: 390, loss: 0.0005535522941499949
step: 400, loss: 0.002113982103765011
step: 410, loss: 8.988850458990782e-05
step: 420, loss: 0.0004310229269322008
step: 430, loss: 4.468236875254661e-05
step: 440, loss: 0.0007574218907393515
step: 450, loss: 2.0637366105802357e-05
step: 460, loss: 0.03915656358003616
step: 470, loss: 2.6087187507073395e-05
step: 480, loss: 6.0568014305317774e-05
step: 490, loss: 0.0003756466321647167
step: 500, loss: 0.00042374030454084277
step: 510, loss: 0.007235510740429163
step: 520, loss: 0.00040786294266581535
step: 530, loss: 0.0022848923690617085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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 14: dev_f1=0.9579831932773109, 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: 3.206804831279442e-05
step: 10, loss: 0.0009406437166035175
step: 20, loss: 0.0012260668445378542
step: 30, loss: 0.0022809007205069065
step: 40, loss: 4.8118799895746633e-05
step: 50, loss: 0.0006089585367590189
step: 60, loss: 0.00038843421498313546
step: 70, loss: 0.0004397130396682769
step: 80, loss: 0.0022713919170200825
step: 90, loss: 0.0001330021768808365
step: 100, loss: 1.2199679076729808e-05
step: 110, loss: 0.004788205958902836
step: 120, loss: 3.961972106480971e-05
step: 130, loss: 8.685459033586085e-05
step: 140, loss: 0.007710837293416262

```

```

step: 150, loss: 0.00017034723714459687
step: 160, loss: 3.158066829200834e-05
step: 170, loss: 5.218040314503014e-05
step: 180, loss: 0.00011643643665593117
step: 190, loss: 0.0020173003431409597
step: 200, loss: 0.0019193398766219616
step: 210, loss: 8.784991223365068e-05
step: 220, loss: 5.9181227697990835e-05
step: 230, loss: 0.0002060584374703467
step: 240, loss: 7.096985791577026e-05
step: 250, loss: 1.9303368389955722e-05
step: 260, loss: 1.8260634533362463e-05
step: 270, loss: 0.0001216508899233304
step: 280, loss: 8.972160867415369e-05
step: 290, loss: 3.092055339948274e-05
step: 300, loss: 4.395133146317676e-05
step: 310, loss: 0.017179690301418304
step: 320, loss: 0.0002502924471627921
step: 330, loss: 3.4091921406798065e-05
step: 340, loss: 0.0004058389749843627
step: 350, loss: 0.0003808413166552782
step: 360, loss: 3.411249053897336e-05
step: 370, loss: 0.00014978715626057237
step: 380, loss: 0.00015477382112294436
step: 390, loss: 0.0005028555751778185
step: 400, loss: 0.0010401514591649175
step: 410, loss: 3.313735214760527e-05
step: 420, loss: 2.6816336685442366e-05
step: 430, loss: 1.1402750715205912e-05
step: 440, loss: 3.0390887332032435e-05
step: 450, loss: 0.0025632278993725777
step: 460, loss: 0.0009100340539589524
step: 470, loss: 5.46757728443481e-05
step: 480, loss: 0.0025303263682872057
step: 490, loss: 3.6608205846278e-05
step: 500, loss: 0.0006698603392578661
step: 510, loss: 2.3702814360149205e-05
step: 520, loss: 3.5796849260805175e-05
step: 530, loss: 0.00019641824474092573
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.958100558659218, f1=0.9455909943714823,
best_f1=0.9455909943714823

```

```
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 [00:34, 168.84it/s]
load_f1 = 0.9596244131455399
real_f1 = 0.9558270676691729
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, 141.51it/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: 0.43662726879119873
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.2745098039215686, f1=0.2692307692307693,
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: 0.41958948969841003
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.3111111111111111, f1=0.3146067415730337,
best_f1=0.3146067415730337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48124608397483826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3943661971830986, f1=0.37681159420289856,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22823600471019745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7272727272727273, f1=0.6956521739130435,
best_f1=0.6956521739130435
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28440961241722107
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9285714285714286, 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: 0.19405347108840942

```

```

/usr/local/lib/python3.7/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=0.9285714285714286, f1=0.8, 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: 0.3107505440711975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.8888888888888889, f1=0.7857142857142857,
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: 0.17370839416980743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.8571428571428571, f1=0.8, 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: 0.043865855783224106
/usr/local/lib/python3.7/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=0.8235294117647058, f1=0.8387096774193549,
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: 0.04406575858592987
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.8484848484848484, f1=0.8125000000000001,
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: 0.017634354531764984
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.8484848484848484, f1=0.8125000000000001,
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: 0.00489927688613534
/usr/local/lib/python3.7/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=0.9032258064516129, f1=0.896551724137931,
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: 0.0021197726018726826
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.8484848484848484, f1=0.8666666666666666,
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.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 0, loss: 0.0691436380147934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.8484848484848484, f1=0.8666666666666666,
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: 0.002883859444409609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8484848484848484, f1=0.8666666666666666,
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, 84724.01it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7999999999999999
real_f1 = 0.8125000000000001
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 143.03it/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, 392kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 666kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 404kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 65.8MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

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

```

```

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

```

```

step: 0, loss: 0.5758377313613892
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.47098079323768616
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: 0.48901116847991943
step: 30, loss: 0.30972322821617126
step: 40, loss: 0.17504341900348663
step: 50, loss: 0.255982369184494
step: 60, loss: 0.3049647808074951
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 70, loss: 0.05186089128255844
step: 80, loss: 0.07618017494678497
step: 90, loss: 0.07906665652990341
step: 100, loss: 0.03580288589000702
step: 110, loss: 0.13003745675086975
step: 120, loss: 0.005091493483632803
step: 130, loss: 0.02344907633960247
step: 140, loss: 0.02700873650610447
step: 150, loss: 0.19413535296916962
step: 160, loss: 0.043142933398485184
step: 170, loss: 0.07324842363595963
step: 180, loss: 0.038492415100336075
step: 190, loss: 0.01746637001633644
step: 200, loss: 0.028549140319228172
step: 210, loss: 0.04774976149201393
step: 220, loss: 0.04251304268836975
step: 230, loss: 0.001253985334187746
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9743016759776536, f1=0.9763779527559054,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004305928945541382
step: 10, loss: 0.09080596268177032
step: 20, loss: 0.010567913763225079
step: 30, loss: 0.020884113386273384
step: 40, loss: 0.13125595450401306
step: 50, loss: 0.0033360833767801523
step: 60, loss: 0.006295390892773867
step: 70, loss: 0.015645721927285194

```

```

step: 80, loss: 0.018123559653759003
step: 90, loss: 0.01731322705745697
step: 100, loss: 0.008217084221541882
step: 110, loss: 0.009786243550479412
step: 120, loss: 0.021264206618070602
step: 130, loss: 0.008626447059214115
step: 140, loss: 0.004603804089128971
step: 150, loss: 0.08193877339363098
step: 160, loss: 0.006297116167843342
step: 170, loss: 0.0014683958142995834
step: 180, loss: 0.012179318815469742
step: 190, loss: 0.004457670729607344
step: 200, loss: 0.004753380082547665
step: 210, loss: 0.06118406355381012
step: 220, loss: 0.0012549952371045947
step: 230, loss: 0.0009021625737659633
/usr/local/lib/python3.7/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=0.9723145071982282, f1=0.9731543624161074,
best_f1=0.9763779527559054
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025111090391874313
step: 10, loss: 0.009761382825672626
step: 20, loss: 0.01521588396281004
step: 30, loss: 0.0013446389930322766
step: 40, loss: 0.02491954155266285
step: 50, loss: 0.03838392347097397
step: 60, loss: 0.008853025734424591
step: 70, loss: 0.0015748200239613652
step: 80, loss: 0.0010986400302499533
step: 90, loss: 0.001141043147072196
step: 100, loss: 0.05788126960396767
step: 110, loss: 0.014695645309984684
step: 120, loss: 0.0030849548056721687
step: 130, loss: 0.0020846764091402292
step: 140, loss: 0.0009895454859361053
step: 150, loss: 0.047400444746017456
step: 160, loss: 0.0017474518390372396
step: 170, loss: 0.0013168331934139132
step: 180, loss: 0.005645287223160267
step: 190, loss: 0.020187148824334145
step: 200, loss: 0.0036325647961348295
step: 210, loss: 0.0016380837187170982
step: 220, loss: 0.014428124763071537
step: 230, loss: 0.024243341758847237

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9773755656108598, f1=0.9806598407281, best_f1=0.9806598407281
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009227320551872253
step: 10, loss: 0.0024220531340688467
step: 20, loss: 0.0011417719069868326
step: 30, loss: 0.0013132799649611115
step: 40, loss: 0.13337953388690948
step: 50, loss: 0.0051137288101017475
step: 60, loss: 0.042783547192811966
step: 70, loss: 0.0026019352953881025
step: 80, loss: 0.0015835869126021862
step: 90, loss: 0.0018344116397202015
step: 100, loss: 0.0008109312038868666
step: 110, loss: 0.0006621906068176031
step: 120, loss: 0.08085814863443375
step: 130, loss: 0.06980997323989868
step: 140, loss: 0.0006993836141191423
step: 150, loss: 0.0009265755652450025
step: 160, loss: 0.001554395304992795
step: 170, loss: 0.0038346967194229364
step: 180, loss: 0.09874078631401062
step: 190, loss: 0.0047235433012247086
step: 200, loss: 0.02181149832904339
step: 210, loss: 0.0004615149518940598
step: 220, loss: 0.00022818485740572214
step: 230, loss: 0.000490564969368279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9842696629213483,
best_f1=0.9842696629213483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002453091728966683

```

```

step: 10, loss: 0.0014753072755411267
step: 20, loss: 0.011154773645102978
step: 30, loss: 0.00046051244135014713
step: 40, loss: 0.0006923145847395062
step: 50, loss: 0.0048199002631008625
step: 60, loss: 0.007317135576158762
step: 70, loss: 0.006364623550325632
step: 80, loss: 0.03308270126581192
step: 90, loss: 0.09711691737174988
step: 100, loss: 0.0007081485819071531
step: 110, loss: 0.004509333986788988
step: 120, loss: 0.005563282873481512
step: 130, loss: 0.0010861027985811234
step: 140, loss: 0.006473686080425978
step: 150, loss: 0.20680682361125946
step: 160, loss: 0.0031124905217438936
step: 170, loss: 0.0028470740653574467
step: 180, loss: 0.0023340352345257998
step: 190, loss: 0.022219887003302574
step: 200, loss: 0.0013331155059859157
step: 210, loss: 0.0029097890947014093
step: 220, loss: 0.0015419076662510633
step: 230, loss: 0.00363937602378428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9820627802690582,
best_f1=0.9820627802690582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014065748546272516
step: 10, loss: 0.001972401514649391
step: 20, loss: 0.001458112383261323
step: 30, loss: 0.0013759295688942075
step: 40, loss: 0.00048299142508767545
step: 50, loss: 0.000548852956853807
step: 60, loss: 0.0024001409765332937
step: 70, loss: 0.045291539281606674
step: 80, loss: 0.012937541119754314
step: 90, loss: 0.007696702145040035
step: 100, loss: 0.0008033060585148633
step: 110, loss: 0.00471057603135705
step: 120, loss: 0.000429987208917737

```

```

step: 130, loss: 0.0005600486183539033
step: 140, loss: 0.0004751006490550935
step: 150, loss: 0.0003873325767926872
step: 160, loss: 0.021234076470136642
step: 170, loss: 0.00021728673891630024
step: 180, loss: 0.0024388181045651436
step: 190, loss: 0.0001909134298330173
step: 200, loss: 0.00035510683665052056
step: 210, loss: 0.0007372883846983314
step: 220, loss: 0.00825434923171997
step: 230, loss: 0.002898558508604765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9820224719101124, f1=0.9852774631936579,
best_f1=0.9820627802690582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017581736901775002
step: 10, loss: 0.0007595367496833205
step: 20, loss: 0.0012590293772518635
step: 30, loss: 0.0004375815042294562
step: 40, loss: 0.0015120964962989092
step: 50, loss: 0.0019009870011359453
step: 60, loss: 0.00020765818771906197
step: 70, loss: 0.0002467595040798187
step: 80, loss: 0.0002882468397729099
step: 90, loss: 0.04662659019231796
step: 100, loss: 0.0003619878552854061
step: 110, loss: 0.0017267564544454217
step: 120, loss: 0.00023737468291074038
step: 130, loss: 0.0003315412614028901
step: 140, loss: 0.0001719274587230757
step: 150, loss: 0.004286446142941713
step: 160, loss: 0.0002543099981267005
step: 170, loss: 0.000698343210387975
step: 180, loss: 0.000983359874226153
step: 190, loss: 0.0011110062478110194
step: 200, loss: 0.008802647702395916
step: 210, loss: 0.006233214400708675
step: 220, loss: 0.0010560002410784364
step: 230, loss: 0.0007557007484138012
/usr/local/lib/python3.7/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=0.9841986455981941, f1=0.9864253393665158,
best_f1=0.9820627802690582

```

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

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

```
step: 0, loss: 0.001954214181751013  
step: 10, loss: 0.009143657051026821  
step: 20, loss: 0.00044879934284836054  
step: 30, loss: 0.00021504978940356523  
step: 40, loss: 0.0003775451041292399  
step: 50, loss: 0.00035866687539964914  
step: 60, loss: 0.00018965262279380113  
step: 70, loss: 6.376445526257157e-05  
step: 80, loss: 0.009236063808202744  
step: 90, loss: 7.12065229890868e-05  
step: 100, loss: 0.00010208867024630308  
step: 110, loss: 0.00013794635015074164  
step: 120, loss: 0.0003408346092328429  
step: 130, loss: 0.0003155085723847151  
step: 140, loss: 7.501975778723136e-05  
step: 150, loss: 0.06262718141078949  
step: 160, loss: 0.0012134482385590672  
step: 170, loss: 0.00012802614946849644  
step: 180, loss: 7.15580172254704e-05  
step: 190, loss: 0.002355394186452031  
step: 200, loss: 0.1039852723479271  
step: 210, loss: 0.005249977111816406  
step: 220, loss: 0.0003153164579998702  
step: 230, loss: 0.00025359998107887805
```

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

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

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

```
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=0.9864864864864865, 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: 0.00028341897996142507  
step: 10, loss: 0.000669014872983098  
step: 20, loss: 0.003147491719573736  
step: 30, loss: 0.00013615447096526623  
step: 40, loss: 0.038497861474752426  
step: 50, loss: 0.0004952721646986902  
step: 60, loss: 0.005253591574728489  
step: 70, loss: 0.07027652114629745  
step: 80, loss: 0.00040036390419118106
```

```

step: 90, loss: 0.052753906697034836
step: 100, loss: 0.00015213093138299882
step: 110, loss: 0.0005532507784664631
step: 120, loss: 0.019985094666481018
step: 130, loss: 0.00027980218874290586
step: 140, loss: 0.00016278892871923745
step: 150, loss: 0.0007892060675658286
step: 160, loss: 0.000405497761676088
step: 170, loss: 0.00010211789049208164
step: 180, loss: 9.605556260794401e-05
step: 190, loss: 9.448247874388471e-05
step: 200, loss: 0.0010478422045707703
step: 210, loss: 0.010675428435206413
step: 220, loss: 0.0005387092242017388
step: 230, loss: 7.303983147721738e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9853107344632768, f1=0.9829351535836178,
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: 0.0001037713882396929
step: 10, loss: 0.00111206725705415
step: 20, loss: 0.0003469129151199013
step: 30, loss: 6.097167351981625e-05
step: 40, loss: 0.00047663014265708625
step: 50, loss: 0.000195961823919788
step: 60, loss: 6.955295248189941e-05
step: 70, loss: 0.015355715528130531
step: 80, loss: 4.824867573915981e-05
step: 90, loss: 6.511880928883329e-05
step: 100, loss: 4.840633482672274e-05
step: 110, loss: 0.00457440409809351
step: 120, loss: 5.7919507526094094e-05
step: 130, loss: 0.001687094452790916
step: 140, loss: 4.6719116653548554e-05
step: 150, loss: 0.0193770844489336
step: 160, loss: 2.82513410638785e-05
step: 170, loss: 0.0001107544667320326
step: 180, loss: 0.010411270894110203
step: 190, loss: 0.006839768495410681
step: 200, loss: 0.00020154660160187632
step: 210, loss: 0.0011351356515660882
step: 220, loss: 0.00015510794764850289
step: 230, loss: 0.0004550949379336089
/usr/local/lib/python3.7/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=0.9844444444444443, f1=0.98, 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: 6.073897748137824e-05
step: 10, loss: 8.506569429300725e-05
step: 20, loss: 0.001076996442861855
step: 30, loss: 9.143684292212129e-05
step: 40, loss: 5.8308760344516486e-05
step: 50, loss: 0.00036439436371438205
step: 60, loss: 0.022191664204001427
step: 70, loss: 0.00030194432474672794
step: 80, loss: 0.00029933382757008076
step: 90, loss: 0.014696684665977955
step: 100, loss: 0.001734809367917478
step: 110, loss: 0.0017091260524466634
step: 120, loss: 0.0010160439414903522
step: 130, loss: 0.0009432896040380001
step: 140, loss: 0.011612108908593655
step: 150, loss: 0.0014980138512328267
step: 160, loss: 0.03933548182249069
step: 170, loss: 0.003031609347090125
step: 180, loss: 5.7589568314142525e-05
step: 190, loss: 0.00015856197569519281
step: 200, loss: 0.0007159541710279882
step: 210, loss: 8.244036143878475e-05
step: 220, loss: 0.009619481861591339
step: 230, loss: 7.970798469614238e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9853768278965129, f1=0.9852440408626559,
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: 8.529453771188855e-05
step: 10, loss: 4.7341301979031414e-05
step: 20, loss: 0.03797011822462082
step: 30, loss: 0.036538638174533844
step: 40, loss: 7.767264469293877e-05
step: 50, loss: 0.0001788252266123891
step: 60, loss: 0.00015318133227992803
step: 70, loss: 8.14661179902032e-05
step: 80, loss: 4.2070903873536736e-05
step: 90, loss: 7.725440082140267e-05

```

```

step: 100, loss: 5.412465179688297e-05
step: 110, loss: 0.0005718799657188356
step: 120, loss: 0.00029889625147916377
step: 130, loss: 4.679071207647212e-05
step: 140, loss: 0.0003361627459526062
step: 150, loss: 5.649275044561364e-05
step: 160, loss: 0.00020295563444960862
step: 170, loss: 0.00023328675888478756
step: 180, loss: 0.00012594218424055725
step: 190, loss: 0.0004186078149359673
step: 200, loss: 4.1765761125134304e-05
step: 210, loss: 0.001241059391759336
step: 220, loss: 0.006472581531852484
step: 230, loss: 0.0009413826046511531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9842696629213483, f1=0.987598647125141,
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: 0.0026892609894275665
step: 10, loss: 3.682647002278827e-05
step: 20, loss: 4.792269101017155e-05
step: 30, loss: 3.896772250300273e-05
step: 40, loss: 0.00010514906898606569
step: 50, loss: 0.0015229607233777642
step: 60, loss: 3.739315434359014e-05
step: 70, loss: 0.0004780297167599201
step: 80, loss: 0.0001333059772150591
step: 90, loss: 5.325655729393475e-05
step: 100, loss: 0.0001042790463543497
step: 110, loss: 0.0009750070166774094
step: 120, loss: 4.8735328164184466e-05
step: 130, loss: 6.76694544381462e-05
step: 140, loss: 3.663676397991367e-05
step: 150, loss: 7.585866842418909e-05
step: 160, loss: 0.008341863751411438
step: 170, loss: 3.864456812152639e-05
step: 180, loss: 0.041068945080041885
step: 190, loss: 4.135279596084729e-05
step: 200, loss: 3.2360137993237004e-05
step: 210, loss: 6.39861827949062e-05
step: 220, loss: 3.667780038085766e-05
step: 230, loss: 4.971379166818224e-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=0.9853768278965129, f1=0.9887133182844244,
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: 3.3984117180807516e-05
step: 10, loss: 3.818265031441115e-05
step: 20, loss: 5.6780023442115635e-05
step: 30, loss: 9.574615251040086e-05
step: 40, loss: 0.0003717669751495123
step: 50, loss: 2.8817868951591663e-05
step: 60, loss: 4.7888534027151763e-05
step: 70, loss: 0.0064963530749082565
step: 80, loss: 3.490057497401722e-05
step: 90, loss: 0.00044598145177587867
step: 100, loss: 5.6362598115811124e-05
step: 110, loss: 3.877869312418625e-05
step: 120, loss: 1.6972189769148827e-05
step: 130, loss: 0.00016311241779476404
step: 140, loss: 7.781918975524604e-05
step: 150, loss: 2.4194761863327585e-05
step: 160, loss: 6.529247912112623e-05
step: 170, loss: 0.0019242365378886461
step: 180, loss: 3.201750587322749e-05
step: 190, loss: 0.015195824205875397
step: 200, loss: 5.101785063743591e-05
step: 210, loss: 4.5884302380727604e-05
step: 220, loss: 6.892038800287992e-05
step: 230, loss: 4.62848947790917e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9853768278965129, f1=0.9853768278965129,
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: 0.019170258194208145
step: 10, loss: 5.0866452511399984e-05
step: 20, loss: 0.00026171840727329254
step: 30, loss: 2.991315705003217e-05
step: 40, loss: 5.620060983346775e-05
step: 50, loss: 3.262421159888618e-05
step: 60, loss: 0.035960182547569275
step: 70, loss: 4.603023626259528e-05
step: 80, loss: 3.634567838162184e-05
step: 90, loss: 4.536631968221627e-05

```



```

step: 100, loss: 2.3572874852106906e-05
step: 110, loss: 0.002786064287647605
step: 120, loss: 0.061168771237134933
step: 130, loss: 0.0002740463532973081
step: 140, loss: 0.01185672264546156
step: 150, loss: 5.121925278217532e-05
step: 160, loss: 0.07319315522909164
step: 170, loss: 2.424680678814184e-05
step: 180, loss: 0.00046588695840910077
step: 190, loss: 7.389130041701719e-05
step: 200, loss: 0.002069292590022087
step: 210, loss: 0.03606012091040611
step: 220, loss: 4.41798911197111e-05
step: 230, loss: 6.7070999648422e-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=0.9843749999999999, f1=0.9810479375696767,
best_f1=0.9852774631936579
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

    warnings.warn("An input tensor was not cuda.")
2473it [00:16, 151.67it/s]
load_f1 = 0.9854423292273236
real_f1 = 0.9833147942157954
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, 135.46it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

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

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

An input tensor was not cuda.

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

step: 0, loss: 0.6153661608695984

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: 0.44172120094299316

step: 20, loss: 0.26531296968460083

step: 30, loss: 0.3670230805873871

step: 40, loss: 0.341380774974823

step: 50, loss: 0.4775312840938568

step: 60, loss: 0.15749353170394897

step: 70, loss: 0.17169442772865295

step: 80, loss: 0.07265815138816833

step: 90, loss: 0.3059576153755188

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

step: 100, loss: 0.18922364711761475

step: 110, loss: 0.2277563363313675

step: 120, loss: 0.045523300766944885

step: 130, loss: 0.18234828114509583

step: 140, loss: 0.4933607876300812

step: 150, loss: 0.1264919936656952

step: 160, loss: 0.18747107684612274

step: 170, loss: 0.032287392765283585

step: 180, loss: 0.013294211588799953

step: 190, loss: 0.13286864757537842

```

step: 200, loss: 0.04610661789774895
step: 210, loss: 0.14447954297065735
step: 220, loss: 0.04937649518251419
step: 230, loss: 0.09895627200603485
step: 240, loss: 0.12135427445173264
step: 250, loss: 0.15493686497211456
step: 260, loss: 0.13328567147254944
step: 270, loss: 0.41818615794181824
step: 280, loss: 0.11625272035598755
step: 290, loss: 0.06577500700950623
step: 300, loss: 0.14162908494472504
step: 310, loss: 0.08181478083133698
step: 320, loss: 0.05053752288222313
step: 330, loss: 0.05033710598945618
step: 340, loss: 0.5093269944190979
step: 350, loss: 0.28579914569854736
step: 360, loss: 0.03744648024439812
step: 370, loss: 0.04513855278491974
step: 380, loss: 0.05734121426939964
step: 390, loss: 0.029410067945718765
step: 400, loss: 0.04078352451324463
step: 410, loss: 0.24499566853046417
step: 420, loss: 0.13102751970291138
step: 430, loss: 0.03677905350923538
step: 440, loss: 0.07855140417814255
step: 450, loss: 0.13989146053791046
step: 460, loss: 0.010244457982480526
step: 470, loss: 0.017313215881586075
step: 480, loss: 0.07766520231962204
step: 490, loss: 0.3171131908893585
step: 500, loss: 0.1467871516942978
step: 510, loss: 0.0800335481762886
step: 520, loss: 0.09701080620288849
step: 530, loss: 0.015716051682829857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9419889502762431, f1=0.9384191176470588,
best_f1=0.9384191176470588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014574511907994747
step: 10, loss: 0.08852457255125046

```

step: 20, loss: 0.014568494632840157
step: 30, loss: 0.032992854714393616
step: 40, loss: 0.0364985428750515
step: 50, loss: 0.05304885283112526
step: 60, loss: 0.08190079033374786
step: 70, loss: 0.028166599571704865
step: 80, loss: 0.016767997294664383
step: 90, loss: 0.020176302641630173
step: 100, loss: 0.06649770587682724
step: 110, loss: 0.01296960934996605
step: 120, loss: 0.034565020352602005
step: 130, loss: 0.003723635571077466
step: 140, loss: 0.07124794274568558
step: 150, loss: 0.05557354912161827
step: 160, loss: 0.07413443922996521
step: 170, loss: 0.06201980635523796
step: 180, loss: 0.017686016857624054
step: 190, loss: 0.04457394406199455
step: 200, loss: 0.16283570230007172
step: 210, loss: 0.013962465338408947
step: 220, loss: 0.0015855047386139631
step: 230, loss: 0.19317008554935455
step: 240, loss: 0.03699071332812309
step: 250, loss: 0.03574026748538017
step: 260, loss: 0.15946218371391296
step: 270, loss: 0.026789801195263863
step: 280, loss: 0.14203676581382751
step: 290, loss: 0.05070304125547409
step: 300, loss: 0.0584280788898468
step: 310, loss: 0.10226092487573624
step: 320, loss: 0.10037514567375183
step: 330, loss: 0.0872906893491745
step: 340, loss: 0.12221741676330566
step: 350, loss: 0.0029641424771398306
step: 360, loss: 0.12441878765821457
step: 370, loss: 0.014582756906747818
step: 380, loss: 0.1578456163406372
step: 390, loss: 0.008975415490567684
step: 400, loss: 0.07949118316173553
step: 410, loss: 0.1548091024160385
step: 420, loss: 0.024943437427282333
step: 430, loss: 0.22583840787410736
step: 440, loss: 0.012372428551316261
step: 450, loss: 0.051255978643894196
step: 460, loss: 0.07927743345499039
step: 470, loss: 0.06926930695772171
step: 480, loss: 0.013716749846935272
step: 490, loss: 0.05710185691714287

```

step: 500, loss: 0.017471596598625183
step: 510, loss: 0.016780875623226166
step: 520, loss: 0.37857785820961
step: 530, loss: 0.059633031487464905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.947075208913649, f1=0.9429763560500695,
best_f1=0.9429763560500695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09135536849498749
step: 10, loss: 0.06876348704099655
step: 20, loss: 0.010309919714927673
step: 30, loss: 0.03498213365674019
step: 40, loss: 0.036243245005607605
step: 50, loss: 0.0281266700476408
step: 60, loss: 0.02161259576678276
step: 70, loss: 0.004830461461097002
step: 80, loss: 0.013538910076022148
step: 90, loss: 0.17641505599021912
step: 100, loss: 0.058334965258836746
step: 110, loss: 0.028967827558517456
step: 120, loss: 0.1146550178527832
step: 130, loss: 0.11324271559715271
step: 140, loss: 0.0290924571454525
step: 150, loss: 0.05473387613892555
step: 160, loss: 0.04767554998397827
step: 170, loss: 0.005927346181124449
step: 180, loss: 0.014686069451272488
step: 190, loss: 0.0021579598542302847
step: 200, loss: 0.027889536693692207
step: 210, loss: 0.07553955912590027
step: 220, loss: 0.14539997279644012
step: 230, loss: 0.07195257395505905
step: 240, loss: 0.042782366275787354
step: 250, loss: 0.059440407902002335
step: 260, loss: 0.077959343791008
step: 270, loss: 0.009466797113418579
step: 280, loss: 0.004398183431476355
step: 290, loss: 0.02579139731824398
step: 300, loss: 0.11070416122674942
step: 310, loss: 0.047500595450401306

```

```

step: 320, loss: 0.06570102274417877
step: 330, loss: 0.011766362003982067
step: 340, loss: 0.016009746119379997
step: 350, loss: 0.12356845289468765
step: 360, loss: 0.09533947706222534
step: 370, loss: 0.01836773194372654
step: 380, loss: 0.023197995498776436
step: 390, loss: 0.006183985620737076
step: 400, loss: 0.08212967216968536
step: 410, loss: 0.08050872385501862
step: 420, loss: 0.011308173649013042
step: 430, loss: 0.03312861546874046
step: 440, loss: 0.18084284663200378
step: 450, loss: 0.18776650726795197
step: 460, loss: 0.047291114926338196
step: 470, loss: 0.08351436257362366
step: 480, loss: 0.12716995179653168
step: 490, loss: 0.024044763296842575
step: 500, loss: 0.07934790849685669
step: 510, loss: 0.05707436800003052
step: 520, loss: 0.004586374387145042
step: 530, loss: 0.0047463420778512955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9517177344475395, f1=0.9447004608294931,
best_f1=0.9447004608294931
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07048104703426361
step: 10, loss: 0.011471197940409184
step: 20, loss: 0.02980980835855007
step: 30, loss: 0.10536158084869385
step: 40, loss: 0.004149760119616985
step: 50, loss: 0.07333196699619293
step: 60, loss: 0.0025904588401317596
step: 70, loss: 0.03964189067482948
step: 80, loss: 0.027732260525226593
step: 90, loss: 0.041711099445819855
step: 100, loss: 0.010736098513007164
step: 110, loss: 0.11785096675157547
step: 120, loss: 0.014968274161219597
step: 130, loss: 0.07761015743017197

```

```

step: 140, loss: 0.03323344513773918
step: 150, loss: 0.10662733018398285
step: 160, loss: 0.004701593890786171
step: 170, loss: 0.08719416707754135
step: 180, loss: 0.059683166444301605
step: 190, loss: 0.057053521275520325
step: 200, loss: 0.06406974792480469
step: 210, loss: 0.0015827121678739786
step: 220, loss: 0.003894635709002614
step: 230, loss: 0.015105972066521645
step: 240, loss: 0.08148838579654694
step: 250, loss: 0.11260472238063812
step: 260, loss: 0.003038247348740697
step: 270, loss: 0.07533905655145645
step: 280, loss: 0.0037250311579555273
step: 290, loss: 0.022807547822594643
step: 300, loss: 0.003573374357074499
step: 310, loss: 0.04476359486579895
step: 320, loss: 0.027835799381136894
step: 330, loss: 0.012019077315926552
step: 340, loss: 0.006285007111728191
step: 350, loss: 0.109251007437706
step: 360, loss: 0.030311714857816696
step: 370, loss: 0.009585512802004814
step: 380, loss: 0.0040499987080693245
step: 390, loss: 0.0009093537810258567
step: 400, loss: 0.011714594438672066
step: 410, loss: 0.001055857166647911
step: 420, loss: 0.015164511278271675
step: 430, loss: 0.023550033569335938
step: 440, loss: 0.0038439584895968437
step: 450, loss: 0.016025485470891
step: 460, loss: 0.01584663800895214
step: 470, loss: 0.0009559783502481878
step: 480, loss: 0.004112562630325556
step: 490, loss: 0.029720541089773178
step: 500, loss: 0.19771035015583038
step: 510, loss: 0.09028945118188858
step: 520, loss: 0.017650999128818512
step: 530, loss: 0.045179009437561035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9560795191863153, f1=0.9411223551057958,

```



```
best_f1=0.9411223551057958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003102777700405568
step: 10, loss: 0.012450474314391613
step: 20, loss: 0.01013158354908228
step: 30, loss: 0.11853048205375671
step: 40, loss: 0.015153211541473866
step: 50, loss: 0.01939535327255726
step: 60, loss: 0.014214437454938889
step: 70, loss: 0.0030171312391757965
step: 80, loss: 0.0001342433679383248
step: 90, loss: 0.2017364799976349
step: 100, loss: 0.1017775684595108
step: 110, loss: 0.014646099880337715
step: 120, loss: 0.13459423184394836
step: 130, loss: 0.0204495619982481
step: 140, loss: 0.015577034093439579
step: 150, loss: 0.031072132289409637
step: 160, loss: 0.026379888877272606
step: 170, loss: 0.08729350566864014
step: 180, loss: 0.00715355621650815
step: 190, loss: 0.007059299852699041
step: 200, loss: 0.010722615756094456
step: 210, loss: 0.0070613110437989235
step: 220, loss: 0.0033319827634841204
step: 230, loss: 0.002696521347388625
step: 240, loss: 0.002115911105647683
step: 250, loss: 0.07942771166563034
step: 260, loss: 0.010189495049417019
step: 270, loss: 0.0032566466834396124
step: 280, loss: 0.006350214127451181
step: 290, loss: 0.007925720885396004
step: 300, loss: 0.03688844293355942
step: 310, loss: 0.03592744097113609
step: 320, loss: 0.10777484625577927
step: 330, loss: 0.0016030221013352275
step: 340, loss: 0.00109589914791286
step: 350, loss: 0.0011367315892130136
step: 360, loss: 0.00013753035455010831
step: 370, loss: 0.0004194128850940615
step: 380, loss: 0.0003297951479908079
step: 390, loss: 0.0014017409412190318
step: 400, loss: 0.048874933272600174
step: 410, loss: 0.05560450255870819
step: 420, loss: 0.169094517827034
step: 430, loss: 0.004613871220499277
```

```

step: 440, loss: 0.02587740682065487
step: 450, loss: 0.0012384873116388917
step: 460, loss: 0.023963652551174164
step: 470, loss: 0.07277469336986542
step: 480, loss: 0.022249983623623848
step: 490, loss: 0.01838034763932228
step: 500, loss: 0.007091633975505829
step: 510, loss: 0.008966200053691864
step: 520, loss: 0.021555617451667786
step: 530, loss: 0.08666466921567917
/usr/local/lib/python3.7/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=0.9527410207939508, f1=0.9421800947867298,
best_f1=0.9411223551057958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02566787227988243
step: 10, loss: 0.1132340133190155
step: 20, loss: 0.0006848240154795349
step: 30, loss: 0.0006158280302770436
step: 40, loss: 0.00036711464053951204
step: 50, loss: 0.0014475106727331877
step: 60, loss: 0.001237240619957447
step: 70, loss: 0.0027084543835371733
step: 80, loss: 0.013895250856876373
step: 90, loss: 0.0049593448638916016
step: 100, loss: 0.0013382986653596163
step: 110, loss: 0.0018351385369896889
step: 120, loss: 0.05346829071640968
step: 130, loss: 0.001937447115778923
step: 140, loss: 0.011866726912558079
step: 150, loss: 0.006138483062386513
step: 160, loss: 0.11004358530044556
step: 170, loss: 0.00853390246629715
step: 180, loss: 0.003641145071014762
step: 190, loss: 0.037188101559877396
step: 200, loss: 0.10352032631635666
step: 210, loss: 0.004144368693232536
step: 220, loss: 0.011211803182959557
step: 230, loss: 0.08397118747234344
step: 240, loss: 0.004484714008867741
step: 250, loss: 0.009230204857885838
step: 260, loss: 0.002452159533277154
step: 270, loss: 0.003838133765384555
step: 280, loss: 0.029318666085600853
step: 290, loss: 0.0234015304595232

```

```

step: 300, loss: 0.0009585358202457428
step: 310, loss: 0.016686927527189255
step: 320, loss: 7.449771510437131e-05
step: 330, loss: 0.018108930438756943
step: 340, loss: 0.0009288133587688208
step: 350, loss: 0.002846227493137121
step: 360, loss: 0.009619603864848614
step: 370, loss: 0.026190828531980515
step: 380, loss: 0.0001439917687093839
step: 390, loss: 0.00213505607098341
step: 400, loss: 0.006100774742662907
step: 410, loss: 0.00041473377496004105
step: 420, loss: 0.01724027842283249
step: 430, loss: 0.014050750993192196
step: 440, loss: 0.03793008252978325
step: 450, loss: 0.24999505281448364
step: 460, loss: 0.0008345910464413464
step: 470, loss: 0.002447264501824975
step: 480, loss: 0.004481528885662556
step: 490, loss: 0.007893389090895653
step: 500, loss: 0.0007863430073484778
step: 510, loss: 0.20867198705673218
step: 520, loss: 0.0049956850707530975
step: 530, loss: 0.005242170765995979
/usr/local/lib/python3.7/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=0.9518916394208314, f1=0.9443144595226953,
best_f1=0.9411223551057958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006387393921613693
step: 10, loss: 0.004052579868584871
step: 20, loss: 0.0027063535526394844
step: 30, loss: 0.01846785843372345
step: 40, loss: 0.0025960730854421854
step: 50, loss: 0.04878489673137665
step: 60, loss: 0.025024961680173874
step: 70, loss: 0.011777461506426334
step: 80, loss: 0.0006859283894300461
step: 90, loss: 0.0006246439297683537
step: 100, loss: 0.0037128396797925234
step: 110, loss: 0.0009897545678541064
step: 120, loss: 0.026239439845085144
step: 130, loss: 0.001916430308483541
step: 140, loss: 0.0014256332069635391
step: 150, loss: 0.0007791652460582554

```

```

step: 160, loss: 0.00024217097961809486
step: 170, loss: 0.050413984805345535
step: 180, loss: 0.041189275681972504
step: 190, loss: 0.03402004763484001
step: 200, loss: 0.0004500004288274795
step: 210, loss: 0.00887291133403778
step: 220, loss: 0.012707572430372238
step: 230, loss: 0.010557572357356548
step: 240, loss: 0.03079768270254135
step: 250, loss: 0.009953219443559647
step: 260, loss: 0.0002540672430768609
step: 270, loss: 0.003350325394421816
step: 280, loss: 0.09710996598005295
step: 290, loss: 0.00376719213090837
step: 300, loss: 0.00048075109953060746
step: 310, loss: 0.0015311356401070952
step: 320, loss: 0.14546984434127808
step: 330, loss: 0.002057512290775776
step: 340, loss: 0.0006534610292874277
step: 350, loss: 0.0053344652988016605
step: 360, loss: 0.0429578498005867
step: 370, loss: 0.10026136040687561
step: 380, loss: 0.05660398676991463
step: 390, loss: 0.07737047225236893
step: 400, loss: 0.012634897604584694
step: 410, loss: 0.01607655920088291
step: 420, loss: 0.010859484784305096
step: 430, loss: 0.0002939327678177506
step: 440, loss: 0.010800976306200027
step: 450, loss: 0.009757773019373417
step: 460, loss: 0.02129635214805603
step: 470, loss: 0.257131427526474
step: 480, loss: 0.0030086440965533257
step: 490, loss: 0.005977381486445665
step: 500, loss: 0.006664665415883064
step: 510, loss: 0.0015216695610433817
step: 520, loss: 0.0008062576525844634
step: 530, loss: 0.0014706358779221773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9583138173302108, f1=0.9417293233082707,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0007879341719672084
step: 10, loss: 0.0027734905015677214
step: 20, loss: 0.0003132297133561224
step: 30, loss: 0.0002563775051385164
step: 40, loss: 0.00011610651563387364
step: 50, loss: 0.0013119240757077932
step: 60, loss: 0.006927804555743933
step: 70, loss: 0.0031365591567009687
step: 80, loss: 0.0008492705528624356
step: 90, loss: 0.00012072834215359762
step: 100, loss: 0.0006346798036247492
step: 110, loss: 9.01187740964815e-05
step: 120, loss: 0.005060513969510794
step: 130, loss: 4.5858432713430375e-05
step: 140, loss: 0.0016636570217087865
step: 150, loss: 0.0008785005193203688
step: 160, loss: 0.00019140199583489448
step: 170, loss: 0.2981850802898407
step: 180, loss: 0.00041799951577559114
step: 190, loss: 0.00945277139544487
step: 200, loss: 0.06304903328418732
step: 210, loss: 0.05099772661924362
step: 220, loss: 0.00020611852232832462
step: 230, loss: 0.02888570912182331
step: 240, loss: 0.007128092460334301
step: 250, loss: 0.003877633949741721
step: 260, loss: 0.0027141417376697063
step: 270, loss: 0.0046813529916107655
step: 280, loss: 0.006332810502499342
step: 290, loss: 0.0037884353660047054
step: 300, loss: 0.0001384321803925559
step: 310, loss: 0.010039453394711018
step: 320, loss: 0.0004849383549299091
step: 330, loss: 8.984318264992908e-05
step: 340, loss: 0.03793127089738846
step: 350, loss: 5.419025183073245e-05
step: 360, loss: 0.024016043171286583
step: 370, loss: 0.017956212162971497
step: 380, loss: 0.00013191676407586783
step: 390, loss: 0.0005624580080620944
step: 400, loss: 0.0005987569456920028
step: 410, loss: 0.013952679932117462
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 420, loss: 0.0011057406663894653
step: 430, loss: 0.0008359994972124696
step: 440, loss: 0.16061916947364807
```

```

step: 450, loss: 0.00455694692209363
step: 460, loss: 0.0026956747751682997
step: 470, loss: 0.07380226999521255
step: 480, loss: 0.04986744001507759
step: 490, loss: 0.024687692523002625
step: 500, loss: 0.01275002770125866
step: 510, loss: 0.011284269392490387
step: 520, loss: 0.00018102809553965926
step: 530, loss: 0.0012728198198601604
/usr/local/lib/python3.7/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=0.9534347625633932, f1=0.9438822447102115,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033767372951842844
step: 10, loss: 0.025634346529841423
step: 20, loss: 5.3679566917708144e-05
step: 30, loss: 0.10295845568180084
step: 40, loss: 0.004764923360198736
step: 50, loss: 0.002335671102628112
step: 60, loss: 0.0015041150618344545
step: 70, loss: 0.0014761354541406035
step: 80, loss: 0.009762252680957317
step: 90, loss: 0.04290285333991051
step: 100, loss: 0.0016128800343722105
step: 110, loss: 0.01941460743546486
step: 120, loss: 0.018810732290148735
step: 130, loss: 0.008523967117071152
step: 140, loss: 0.00040753098437562585
step: 150, loss: 0.009107553400099277
step: 160, loss: 0.0003342706768307835
step: 170, loss: 0.00950646959245205
step: 180, loss: 0.019143905490636826
step: 190, loss: 0.0012422751169651747
step: 200, loss: 0.0023452250752598047
step: 210, loss: 0.01783192902803421
step: 220, loss: 0.0008272908162325621
step: 230, loss: 0.001106612617149949
step: 240, loss: 0.00013224910071585327
step: 250, loss: 0.007296285126358271
step: 260, loss: 0.0022483014035969973
step: 270, loss: 0.002210411475971341
step: 280, loss: 0.06685605645179749
step: 290, loss: 0.002387278713285923
step: 300, loss: 0.00035048284917138517

```

```

step: 310, loss: 0.017613200470805168
step: 320, loss: 9.081740427063778e-05
step: 330, loss: 0.00032509834272786975
step: 340, loss: 0.11804457753896713
step: 350, loss: 0.11581232398748398
step: 360, loss: 0.04543883353471756
step: 370, loss: 0.0012436866527423263
step: 380, loss: 0.0006136518204584718
step: 390, loss: 0.0003710875753313303
step: 400, loss: 0.02813698723912239
step: 410, loss: 0.0009180662455037236
step: 420, loss: 0.00015560029714833945
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 430, loss: 0.19264838099479675
step: 440, loss: 0.0009402437135577202
step: 450, loss: 0.0017037135548889637
step: 460, loss: 0.033051010221242905
step: 470, loss: 0.02350584603846073
step: 480, loss: 0.0016582917887717485
step: 490, loss: 0.02117251232266426
step: 500, loss: 0.001364786410704255
step: 510, loss: 0.0020024196710437536
step: 520, loss: 0.0044100163504481316
step: 530, loss: 0.06035655364394188
/usr/local/lib/python3.7/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=0.9518072289156626, f1=0.94547134935305,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021213379222899675
step: 10, loss: 0.0009805887239053845
step: 20, loss: 0.0013030776754021645
step: 30, loss: 0.00800575502216816
step: 40, loss: 0.0014238134026527405
step: 50, loss: 0.00022297067334875464
step: 60, loss: 0.008036701008677483
step: 70, loss: 7.543063838966191e-05
step: 80, loss: 0.0005795076140202582
step: 90, loss: 0.0023202686570584774
step: 100, loss: 0.007488987408578396
step: 110, loss: 0.006067476700991392
step: 120, loss: 0.00010432057752041146
step: 130, loss: 0.001988620962947607
step: 140, loss: 0.0005860390374436975
step: 150, loss: 0.002543318085372448

```

```

step: 160, loss: 0.05175492912530899
step: 170, loss: 0.00040233571780845523
step: 180, loss: 0.01704711653292179
step: 190, loss: 0.002122286008670926
step: 200, loss: 0.002588327508419752
step: 210, loss: 0.008216765709221363
step: 220, loss: 0.0017973645590245724
step: 230, loss: 0.00028145129908807576
step: 240, loss: 0.0008293692953884602
step: 250, loss: 0.0003392094513401389
step: 260, loss: 0.003845691680908203
step: 270, loss: 0.0001335779088549316
step: 280, loss: 0.042981524020433426
step: 290, loss: 0.0002997714327648282
step: 300, loss: 0.0005621242453344166
step: 310, loss: 0.057244669646024704
step: 320, loss: 0.0012992656556889415
step: 330, loss: 0.0032158740796148777
step: 340, loss: 0.00039381030364893377
step: 350, loss: 0.00027136935386806726
step: 360, loss: 0.00012570768012665212
step: 370, loss: 0.0013060293858870864
step: 380, loss: 0.0015691617736592889
step: 390, loss: 9.139390022028238e-05
step: 400, loss: 0.00041725137270987034
step: 410, loss: 0.01016294676810503
step: 420, loss: 0.0006534992135129869
step: 430, loss: 0.0008830474689602852
step: 440, loss: 0.0006335992948152125
step: 450, loss: 0.0016804223414510489
step: 460, loss: 0.0011649476364254951
step: 470, loss: 0.0016832725377753377
step: 480, loss: 0.003466083202511072
step: 490, loss: 0.0006557385204359889
step: 500, loss: 0.00201254989951849
step: 510, loss: 0.00020665046758949757
step: 520, loss: 0.0014113332144916058
step: 530, loss: 0.003833887865766883
/usr/local/lib/python3.7/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=0.9541029207232267, f1=0.9472710453283996,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.65604675305076e-05
step: 10, loss: 0.00042734917951747775

```


step: 20, loss: 0.0006190871936269104
step: 30, loss: 0.00010025956726167351
step: 40, loss: 6.031654265825637e-05
step: 50, loss: 2.9149303372832946e-05
step: 60, loss: 5.560055797104724e-05
step: 70, loss: 0.00012195414456073195
step: 80, loss: 0.0005165369948372245
step: 90, loss: 0.0005159482243470848
step: 100, loss: 0.0008885283605195582
step: 110, loss: 0.00010008415119955316
step: 120, loss: 0.0037625632248818874
step: 130, loss: 0.016550440341234207
step: 140, loss: 7.451752753695473e-05
step: 150, loss: 0.0008627693168818951
step: 160, loss: 0.0028501185588538647
step: 170, loss: 0.00045300618512555957
step: 180, loss: 0.00011753010767279193
step: 190, loss: 0.00022517792240250856
step: 200, loss: 8.239644375862554e-05
step: 210, loss: 0.00010185291466768831
step: 220, loss: 0.06139079108834267
step: 230, loss: 0.00019053849973715842
step: 240, loss: 0.0011491056066006422
step: 250, loss: 0.00039185379864647985
step: 260, loss: 0.0022400247398763895
step: 270, loss: 0.0009712171158753335
step: 280, loss: 0.0009328118758276105
step: 290, loss: 0.0006575538427568972
step: 300, loss: 0.007667961996048689
step: 310, loss: 0.0005586973275057971
step: 320, loss: 0.00023784313816577196
step: 330, loss: 2.5539984562783502e-05
step: 340, loss: 0.3829532861709595
step: 350, loss: 0.00044868362601846457
step: 360, loss: 0.001490887487307191
step: 370, loss: 0.0015914409887045622
step: 380, loss: 0.0011296847369521856
step: 390, loss: 0.001417831634171307
step: 400, loss: 0.028459345921874046
step: 410, loss: 0.002870081225410104
step: 420, loss: 0.0026761386543512344
step: 430, loss: 0.0030802390538156033
step: 440, loss: 0.004143545404076576
step: 450, loss: 0.0020216715056449175
step: 460, loss: 0.00019168955623172224
step: 470, loss: 0.0010128532303497195
step: 480, loss: 0.0019854314159601927
step: 490, loss: 0.003059369744732976

```

step: 500, loss: 0.001396101899445057
step: 510, loss: 0.0006825204472988844
step: 520, loss: 0.0023500092793256044
step: 530, loss: 0.05529521405696869
/usr/local/lib/python3.7/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=0.9453860640301318, f1=0.9435370975268316,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046324077993631363
step: 10, loss: 0.0036683515645563602
step: 20, loss: 0.004955056123435497
step: 30, loss: 2.4760533051448874e-05
step: 40, loss: 0.0013204419519752264
step: 50, loss: 3.235620533814654e-05
step: 60, loss: 4.35030524386093e-05
step: 70, loss: 0.003163354005664587
step: 80, loss: 1.209544552693842e-05
step: 90, loss: 0.011350294575095177
step: 100, loss: 0.005107852164655924
step: 110, loss: 0.0025425099302083254
step: 120, loss: 0.00013430191029328853
step: 130, loss: 0.0005453547346405685
step: 140, loss: 0.0004067709669470787
step: 150, loss: 9.364107245346531e-05
step: 160, loss: 9.072808461496606e-05
step: 170, loss: 5.516685268958099e-05
step: 180, loss: 0.0012169082183390856
step: 190, loss: 0.0002747699909377843
step: 200, loss: 0.0016058610053732991
step: 210, loss: 0.0015360326506197453
step: 220, loss: 0.006173653528094292
step: 230, loss: 9.181931091006845e-05
step: 240, loss: 0.008025488816201687
step: 250, loss: 1.2486991181503981e-05
step: 260, loss: 0.0013102799421176314
step: 270, loss: 0.037290070205926895
step: 280, loss: 0.000562880712095648
step: 290, loss: 0.0019422030309215188
step: 300, loss: 0.0011202179593965411
step: 310, loss: 7.195038051577285e-05
step: 320, loss: 0.0003211490693502128
step: 330, loss: 0.0015783982817083597
step: 340, loss: 0.0019261296838521957
step: 350, loss: 4.9113739805761725e-05

```

```

step: 360, loss: 3.298891169833951e-05
step: 370, loss: 4.7364555939566344e-05
step: 380, loss: 0.0004354819539003074
step: 390, loss: 0.0038861287757754326
step: 400, loss: 1.019225237541832e-05
step: 410, loss: 0.00043526102672331035
step: 420, loss: 0.001230540219694376
step: 430, loss: 0.00035731293610297143
step: 440, loss: 0.0028491620905697346
step: 450, loss: 1.2654418242163956e-05
step: 460, loss: 1.3112541637383401e-05
step: 470, loss: 0.017993299290537834
step: 480, loss: 0.002803827403113246
step: 490, loss: 0.0015446122270077467
step: 500, loss: 0.00031610787846148014
step: 510, loss: 0.019497569650411606
step: 520, loss: 0.005226809997111559
step: 530, loss: 0.0013261158019304276
/usr/local/lib/python3.7/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=0.9534016775396085, f1=0.9461966604823748,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002067547757178545
step: 10, loss: 9.947256330633536e-05
step: 20, loss: 0.0003754321369342506
step: 30, loss: 4.452863504411653e-05
step: 40, loss: 0.0006791417254135013
step: 50, loss: 0.002644440857693553
step: 60, loss: 0.011227764189243317
step: 70, loss: 0.014580207876861095
step: 80, loss: 0.006933648604899645
step: 90, loss: 0.0003818051191046834
step: 100, loss: 0.004721158184111118
step: 110, loss: 0.0013517123879864812
step: 120, loss: 0.0006855488754808903
step: 130, loss: 0.000997569877654314
step: 140, loss: 0.0007509315037168562
step: 150, loss: 0.001907547004520893
step: 160, loss: 0.00995780061930418
step: 170, loss: 0.005103427451103926
step: 180, loss: 0.0007710533100180328
step: 190, loss: 0.0002711731649469584
step: 200, loss: 0.00043723397538997233
step: 210, loss: 0.00010516280599404126

```

```

step: 220, loss: 0.00023633890668861568
step: 230, loss: 0.012872177176177502
step: 240, loss: 0.0005981817375868559
step: 250, loss: 0.0005884030251763761
step: 260, loss: 0.001047450234182179
step: 270, loss: 0.03603619337081909
step: 280, loss: 0.01989574357867241
step: 290, loss: 0.00019641011022031307
step: 300, loss: 0.0009136603912338614
step: 310, loss: 0.002797322813421488
step: 320, loss: 0.00022462934430222958
step: 330, loss: 0.0004863250069320202
step: 340, loss: 0.003079955466091633
step: 350, loss: 0.0001182737250928767
step: 360, loss: 0.0034217769280076027
step: 370, loss: 0.016352621838450432
step: 380, loss: 0.00048186304047703743
step: 390, loss: 0.0014798571355640888
step: 400, loss: 0.0005000795936211944
step: 410, loss: 0.0006062659085728228
step: 420, loss: 0.00044284301111474633
step: 430, loss: 0.00036271687713451684
step: 440, loss: 4.849258402828127e-05
step: 450, loss: 0.07471774518489838
step: 460, loss: 0.016751276329159737
step: 470, loss: 0.0030875166412442923
step: 480, loss: 1.9468096070340835e-05
step: 490, loss: 7.055843889247626e-05
step: 500, loss: 0.00031016673892736435
step: 510, loss: 8.187827916117385e-05
step: 520, loss: 0.03757556900382042
step: 530, loss: 8.29724085633643e-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=0.9530075187969925, f1=0.9460853258321612,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005845892592333257
step: 10, loss: 0.16723594069480896
step: 20, loss: 0.008310345932841301
step: 30, loss: 0.0007091017905622721
step: 40, loss: 0.00022622337564826012
step: 50, loss: 0.0028724023140966892
step: 60, loss: 0.0010952462907880545
step: 70, loss: 0.00235474924556911

```

step: 80, loss: 0.0008613725076429546
step: 90, loss: 3.2356783776776865e-05
step: 100, loss: 0.001750017050653696
step: 110, loss: 0.006880949251353741
step: 120, loss: 6.485700578195974e-05
step: 130, loss: 0.00037079720641486347
step: 140, loss: 0.008077431470155716
step: 150, loss: 0.0014818933559581637
step: 160, loss: 0.00010305898467777297
step: 170, loss: 0.003617688314989209
step: 180, loss: 0.00016445199435111135
step: 190, loss: 0.0001681274879956618
step: 200, loss: 0.021290984004735947
step: 210, loss: 0.002103836741298437
step: 220, loss: 5.2951141697121784e-05
step: 230, loss: 0.00010008856770582497
step: 240, loss: 0.05097193270921707
step: 250, loss: 0.0017081683035939932
step: 260, loss: 0.0003756181977223605
step: 270, loss: 0.004088377580046654
step: 280, loss: 7.028357504168525e-05
step: 290, loss: 0.00014086069131735712
step: 300, loss: 7.235679368022829e-05
step: 310, loss: 0.0001181244952022098
step: 320, loss: 2.2886479200678878e-05
step: 330, loss: 0.0020598776172846556
step: 340, loss: 0.00042369982111267745
step: 350, loss: 4.531748709268868e-05
step: 360, loss: 0.0017028548754751682
step: 370, loss: 0.08658900111913681
step: 380, loss: 0.010511419735848904
step: 390, loss: 0.00022554895258508623
step: 400, loss: 0.0017517973901703954
step: 410, loss: 0.00021233443112578243
step: 420, loss: 0.00242912326939404
step: 430, loss: 0.0015676453476771712
step: 440, loss: 0.0587652288377285
step: 450, loss: 8.38759369798936e-05
step: 460, loss: 0.048418886959552765
step: 470, loss: 0.0001105812334571965
step: 480, loss: 0.0001324364566244185
step: 490, loss: 0.0014767993707209826
step: 500, loss: 0.004245541989803314
step: 510, loss: 0.022935783490538597
step: 520, loss: 0.0008700334001332521
step: 530, loss: 0.0003749170573428273
/usr/local/lib/python3.7/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=0.9523809523809524, f1=0.9451162790697675,
best_f1=0.9417293233082707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.593549263314344e-05
step: 10, loss: 0.0014788435073569417
step: 20, loss: 0.0008186138584278524
step: 30, loss: 0.00414033979177475
step: 40, loss: 4.9432132072979584e-05
step: 50, loss: 0.10679762065410614
step: 60, loss: 0.00015938404249027371
step: 70, loss: 4.419747710926458e-05
step: 80, loss: 0.002973671304062009
step: 90, loss: 0.00044978666119277477
step: 100, loss: 0.00021318548533599824
step: 110, loss: 0.12133430689573288
step: 120, loss: 0.00011563585576368496
step: 130, loss: 0.0012629467528313398
step: 140, loss: 2.9786297091050074e-05
step: 150, loss: 3.1820170988794416e-05
step: 160, loss: 6.296725041465834e-05
step: 170, loss: 0.00022595500922761858
step: 180, loss: 0.002486372133716941
step: 190, loss: 0.0038893879391252995
step: 200, loss: 0.0009405421442352235
step: 210, loss: 0.0003968336386606097
step: 220, loss: 0.0003774306969717145
step: 230, loss: 0.0003523436316754669
step: 240, loss: 0.003437321400269866
step: 250, loss: 0.00030725522083230317
step: 260, loss: 5.380709990276955e-05
step: 270, loss: 2.681346086319536e-05
step: 280, loss: 2.116303403454367e-05
step: 290, loss: 0.0021502310410141945
step: 300, loss: 0.0008215512498281896
step: 310, loss: 0.03259345889091492
step: 320, loss: 0.0006657131598331034
step: 330, loss: 8.164224709616974e-05
step: 340, loss: 0.0003394647501409054
step: 350, loss: 0.009767135605216026
step: 360, loss: 0.00013448757817968726
step: 370, loss: 0.00473667960613966
step: 380, loss: 0.0003009773208759725
step: 390, loss: 0.00889122486114502
step: 400, loss: 0.02946045994758606
step: 410, loss: 0.0007818790036253631
```

```

step: 420, loss: 0.00034522576606832445
step: 430, loss: 0.00011313310824334621
step: 440, loss: 0.005595327354967594
step: 450, loss: 0.0013451979029923677
step: 460, loss: 0.0024443878792226315
step: 470, loss: 2.4180286345654167e-05
step: 480, loss: 0.002035069977864623
step: 490, loss: 0.00023604884336236864
step: 500, loss: 0.0011107288300991058
step: 510, loss: 0.00018021524010691792
step: 520, loss: 0.0003704004629980773
step: 530, loss: 8.147628250299022e-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=0.9503745318352059, f1=0.9454036397573495,
best_f1=0.9417293233082707
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 [00:33, 171.72it/s]

```

```
load_f1 = 0.9579831932773109
real_f1 = 0.9548206800186306
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.28it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly because apex was installed without --cuda_ext --cpp_ext. Using Python fallback. Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.5200115442276001

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

step: 10, loss: 0.35974815487861633

step: 20, loss: 0.4261551797389984

step: 30, loss: 0.3772832155227661

step: 40, loss: 0.38120681047439575

step: 50, loss: 0.46585822105407715

step: 60, loss: 0.4916466176509857

step: 70, loss: 0.3111608624458313

step: 80, loss: 0.3924912214279175

step: 90, loss: 0.32271289825439453

step: 100, loss: 0.2738018333911896

step: 110, loss: 0.24340660870075226

step: 120, loss: 0.5142697691917419

step: 130, loss: 0.23907452821731567

```

step: 140, loss: 0.4749037027359009
step: 150, loss: 0.403715580701828
step: 160, loss: 0.5181036591529846
step: 170, loss: 0.24031329154968262
step: 180, loss: 0.3980228006839752
step: 190, loss: 0.6303042769432068
step: 200, loss: 0.38058462738990784
step: 210, loss: 0.5020630955696106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.3703288435935974
step: 10, loss: 0.19111622869968414
step: 20, loss: 0.5169988870620728
step: 30, loss: 0.48461493849754333
step: 40, loss: 0.4604407250881195
step: 50, loss: 0.23497363924980164
step: 60, loss: 0.32293280959129333
step: 70, loss: 0.42777353525161743
step: 80, loss: 0.30828890204429626
step: 90, loss: 0.37507787346839905
step: 100, loss: 0.514609158039093
step: 110, loss: 0.37934979796409607
step: 120, loss: 0.23470288515090942
step: 130, loss: 0.17402291297912598
step: 140, loss: 0.23285672068595886
step: 150, loss: 0.44041019678115845
step: 160, loss: 0.1661197990179062
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.63548344373703
step: 180, loss: 0.3202177882194519
step: 190, loss: 0.30531638860702515
step: 200, loss: 0.15158911049365997
step: 210, loss: 0.3048604726791382
/usr/local/lib/python3.7/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: 0.25517335534095764
step: 10, loss: 0.2507563829421997
step: 20, loss: 0.4960942566394806
step: 30, loss: 0.2501380443572998
step: 40, loss: 0.4545533359050751
step: 50, loss: 0.47746405005455017
step: 60, loss: 0.5107579231262207
step: 70, loss: 0.2071477174758911
step: 80, loss: 0.47510331869125366
step: 90, loss: 0.23934023082256317
step: 100, loss: 0.36673611402511597
step: 110, loss: 0.23077066242694855
step: 120, loss: 0.22643128037452698
step: 130, loss: 0.1568623036146164
step: 140, loss: 0.37185463309288025
step: 150, loss: 0.32092320919036865
step: 160, loss: 0.20692217350006104
step: 170, loss: 0.36875540018081665
step: 180, loss: 0.2497713565826416
step: 190, loss: 0.16217531263828278
step: 200, loss: 0.23670396208763123
step: 210, loss: 0.2565300464630127
/usr/local/lib/python3.7/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: 0.32363811135292053
step: 10, loss: 0.3224085569381714
step: 20, loss: 0.3050297498703003
step: 30, loss: 0.23477934300899506
step: 40, loss: 0.23131202161312103
step: 50, loss: 0.24309787154197693
step: 60, loss: 0.5037844777107239
step: 70, loss: 0.2501809298992157
step: 80, loss: 0.24635037779808044
step: 90, loss: 0.46045053005218506
step: 100, loss: 0.39089781045913696
step: 110, loss: 0.5777586698532104
step: 120, loss: 0.3745858669281006
step: 130, loss: 0.6868493556976318
step: 140, loss: 0.48048919439315796

```

```

step: 150, loss: 0.39623671770095825
step: 160, loss: 0.3102531433105469
step: 170, loss: 0.16946718096733093
step: 180, loss: 0.07755548506975174
step: 190, loss: 0.17816273868083954
step: 200, loss: 0.3087047338485718
step: 210, loss: 0.3988575041294098
/usr/local/lib/python3.7/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: 0.3859573006629944
step: 10, loss: 0.31360822916030884
step: 20, loss: 0.31108081340789795
step: 30, loss: 0.23445676267147064
step: 40, loss: 0.436363160610199
step: 50, loss: 0.3822958171367645
step: 60, loss: 0.3814581632614136
step: 70, loss: 0.24200215935707092
step: 80, loss: 0.453311949968338
step: 90, loss: 0.4266911745071411
step: 100, loss: 0.24451400339603424
step: 110, loss: 0.16929443180561066
step: 120, loss: 0.24005888402462006
step: 130, loss: 0.3074559271335602
step: 140, loss: 0.5358681678771973
step: 150, loss: 0.31391441822052
step: 160, loss: 0.25128471851348877
step: 170, loss: 0.3812582194805145
step: 180, loss: 0.2384275197982788
step: 190, loss: 0.5265268087387085
step: 200, loss: 0.3831312358379364
step: 210, loss: 0.24154460430145264
/usr/local/lib/python3.7/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: 0.1655685305595398
step: 10, loss: 0.3677610754966736
step: 20, loss: 0.3083846867084503

```

```

step: 30, loss: 0.24119128286838531
step: 40, loss: 0.30740097165107727
step: 50, loss: 0.6006051301956177
step: 60, loss: 0.31598538160324097
step: 70, loss: 0.44655677676200867
step: 80, loss: 0.31642237305641174
step: 90, loss: 0.3061197102069855
step: 100, loss: 0.3051886558532715
step: 110, loss: 0.3767683506011963
step: 120, loss: 0.24544622004032135
step: 130, loss: 0.2372761070728302
step: 140, loss: 0.3788208067417145
step: 150, loss: 0.1925632804632187
step: 160, loss: 0.16436554491519928
step: 170, loss: 0.4579062759876251
step: 180, loss: 0.3758019506931305
step: 190, loss: 0.44978800415992737
step: 200, loss: 0.311418741941452
step: 210, loss: 0.3869072496891022
/usr/local/lib/python3.7/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: 0.4602774679660797
step: 10, loss: 0.31821104884147644
step: 20, loss: 0.44311895966529846
step: 30, loss: 0.24021360278129578
step: 40, loss: 0.2431579977273941
step: 50, loss: 0.38500073552131653
step: 60, loss: 0.30143195390701294
step: 70, loss: 0.25255122780799866
step: 80, loss: 0.5003869533538818
step: 90, loss: 0.3849400579929352
step: 100, loss: 0.45513394474983215
step: 110, loss: 0.3142758905887604
step: 120, loss: 0.30307355523109436
step: 130, loss: 0.4680027961730957
step: 140, loss: 0.3131594955921173
step: 150, loss: 0.31385666131973267
step: 160, loss: 0.5561776161193848
step: 170, loss: 0.6206589937210083
step: 180, loss: 0.2519952058792114
step: 190, loss: 0.2430519312620163
step: 200, loss: 0.3173801302909851

```

```

step: 210, loss: 0.3207732141017914
/usr/local/lib/python3.7/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: 0.6054751873016357
step: 10, loss: 0.3166385591030121
step: 20, loss: 0.25740835070610046
step: 30, loss: 0.25056612491607666
step: 40, loss: 0.24035939574241638
step: 50, loss: 0.1670786589384079
step: 60, loss: 0.17102967202663422
step: 70, loss: 0.4638727605342865
step: 80, loss: 0.3200351595878601
step: 90, loss: 0.3779195547103882
step: 100, loss: 0.62522953748703
step: 110, loss: 0.3910450339317322
step: 120, loss: 0.30894699692726135
step: 130, loss: 0.17244090139865875
step: 140, loss: 0.3111981749534607
step: 150, loss: 0.4645046293735504
step: 160, loss: 0.45172449946403503
step: 170, loss: 0.521863579750061
step: 180, loss: 0.3075255751609802
step: 190, loss: 0.22887884080410004
step: 200, loss: 0.45747262239456177
step: 210, loss: 0.3792479336261749
/usr/local/lib/python3.7/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: 0.5917030572891235
step: 10, loss: 0.26919835805892944
step: 20, loss: 0.37844714522361755
step: 30, loss: 0.17302221059799194
step: 40, loss: 0.16720221936702728
step: 50, loss: 0.4579366147518158
step: 60, loss: 0.17848047614097595
step: 70, loss: 0.3897424638271332
step: 80, loss: 0.3122943639755249

```

```

step: 90, loss: 0.3108953535556793
step: 100, loss: 0.4660419523715973
step: 110, loss: 0.5386745929718018
step: 120, loss: 0.38564640283584595
step: 130, loss: 0.24955077469348907
step: 140, loss: 0.5926229357719421
step: 150, loss: 0.11936823278665543
step: 160, loss: 0.3840784430503845
step: 170, loss: 0.31056034564971924
step: 180, loss: 0.4497436583042145
step: 190, loss: 0.3166910409927368
step: 200, loss: 0.3046651780605316
step: 210, loss: 0.31579411029815674
/usr/local/lib/python3.7/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: 0.3198948800563812
step: 10, loss: 0.372403085231781
step: 20, loss: 0.24947215616703033
step: 30, loss: 0.1749790757894516
step: 40, loss: 0.3098183870315552
step: 50, loss: 0.5098845362663269
step: 60, loss: 0.24298320710659027
step: 70, loss: 0.38327789306640625
step: 80, loss: 0.16112485527992249
step: 90, loss: 0.3909149765968323
step: 100, loss: 0.4674236476421356
step: 110, loss: 0.17451316118240356
step: 120, loss: 0.4454503059387207
step: 130, loss: 0.24255657196044922
step: 140, loss: 0.35945937037467957
step: 150, loss: 0.25074079632759094
step: 160, loss: 0.37707698345184326
step: 170, loss: 0.24657920002937317
step: 180, loss: 0.30908626317977905
step: 190, loss: 0.24477989971637726
step: 200, loss: 0.6556708812713623
step: 210, loss: 0.25573715567588806
/usr/local/lib/python3.7/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: 0.3085307776927948
step: 10, loss: 0.315737783908844
step: 20, loss: 0.37505847215652466
step: 30, loss: 0.1725032776594162
step: 40, loss: 0.4573073387145996
step: 50, loss: 0.4422983229160309
step: 60, loss: 0.443647176027298
step: 70, loss: 0.17073634266853333
step: 80, loss: 0.5230012536048889
step: 90, loss: 0.38066574931144714
step: 100, loss: 0.5135971903800964
step: 110, loss: 0.44302117824554443
step: 120, loss: 0.436780720949173
step: 130, loss: 0.24414679408073425
step: 140, loss: 0.305787593126297
step: 150, loss: 0.31101495027542114
step: 160, loss: 0.16775810718536377
step: 170, loss: 0.31204840540885925
step: 180, loss: 0.243128702044487
step: 190, loss: 0.4471651613712311
step: 200, loss: 0.15820488333702087
step: 210, loss: 0.3630239963531494
/usr/local/lib/python3.7/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: 0.25427693128585815
step: 10, loss: 0.30999046564102173
step: 20, loss: 0.44753748178482056
step: 30, loss: 0.2924090027809143
step: 40, loss: 0.17469921708106995
step: 50, loss: 0.45601752400398254
step: 60, loss: 0.10737062990665436
step: 70, loss: 0.4466947019100189
step: 80, loss: 0.30596816539764404
step: 90, loss: 0.38638752698898315
step: 100, loss: 0.12326637655496597
step: 110, loss: 0.253410667181015
step: 120, loss: 0.2473495602607727
step: 130, loss: 0.4489307999610901
step: 140, loss: 0.3843473792076111

```



```

step: 150, loss: 0.1019800677895546
step: 160, loss: 0.2478903830051422
step: 170, loss: 0.24127180874347687
step: 180, loss: 0.31043222546577454
step: 190, loss: 0.36990123987197876
step: 200, loss: 0.18445120751857758
step: 210, loss: 0.3110105097293854
/usr/local/lib/python3.7/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: 0.3162502944469452
step: 10, loss: 0.24128788709640503
step: 20, loss: 0.3162675201892853
step: 30, loss: 0.24568068981170654
step: 40, loss: 0.2448110431432724
step: 50, loss: 0.24246445298194885
step: 60, loss: 0.574219286441803
step: 70, loss: 0.4467727243900299
step: 80, loss: 0.3779783546924591
step: 90, loss: 0.24189071357250214
step: 100, loss: 0.3799103796482086
step: 110, loss: 0.24484038352966309
step: 120, loss: 0.31063854694366455
step: 130, loss: 0.30948176980018616
step: 140, loss: 0.164278045296669
step: 150, loss: 0.315461128950119
step: 160, loss: 0.5285632610321045
step: 170, loss: 0.23941002786159515
step: 180, loss: 0.23979558050632477
step: 190, loss: 0.2416006624698639
step: 200, loss: 0.3813256621360779
step: 210, loss: 0.3715396225452423
/usr/local/lib/python3.7/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: 0.24574637413024902
step: 10, loss: 0.37704312801361084
step: 20, loss: 0.4985025227069855

```

```

step: 30, loss: 0.24424338340759277
step: 40, loss: 0.3092271685600281
step: 50, loss: 0.25102466344833374
step: 60, loss: 0.3874599039554596
step: 70, loss: 0.25471964478492737
step: 80, loss: 0.31542524695396423
step: 90, loss: 0.1801607757806778
step: 100, loss: 0.2540598213672638
step: 110, loss: 0.4569617807865143
step: 120, loss: 0.24380099773406982
step: 130, loss: 0.31481096148490906
step: 140, loss: 0.3124077022075653
step: 150, loss: 0.31337082386016846
step: 160, loss: 0.17192456126213074
step: 170, loss: 0.44984927773475647
step: 180, loss: 0.2402227818965912
step: 190, loss: 0.380230188369751
step: 200, loss: 0.1678464710712433
step: 210, loss: 0.4553298056125641
/usr/local/lib/python3.7/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: 0.2460293173789978
step: 10, loss: 0.24721381068229675
step: 20, loss: 0.5246127843856812
step: 30, loss: 0.24565979838371277
step: 40, loss: 0.38639959692955017
step: 50, loss: 0.17285309731960297
step: 60, loss: 0.3150438666343689
step: 70, loss: 0.44908344745635986
step: 80, loss: 0.24475061893463135
step: 90, loss: 0.44999784231185913
step: 100, loss: 0.3020312488079071
step: 110, loss: 0.317268431186676
step: 120, loss: 0.4502562880516052
step: 130, loss: 0.17127138376235962
step: 140, loss: 0.31300580501556396
step: 150, loss: 0.5255082845687866
step: 160, loss: 0.37744778394699097
step: 170, loss: 0.38361647725105286
step: 180, loss: 0.31597450375556946
step: 190, loss: 0.3144724667072296
step: 200, loss: 0.3126302659511566

```

```

step: 210, loss: 0.4478510618209839
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:09, 235.32it/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 [00:31, 140.87it/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: 0.4689573347568512
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.46633413434028625
step: 20, loss: 0.24646641314029694
step: 30, loss: 0.41700807213783264
step: 40, loss: 0.20724768936634064
step: 50, loss: 0.3158736526966095
step: 60, loss: 0.42093634605407715
step: 70, loss: 0.44951131939888
step: 80, loss: 0.19532454013824463
step: 90, loss: 0.29565146565437317
step: 100, loss: 0.4019623398780823
step: 110, loss: 0.21738123893737793
step: 120, loss: 0.32855063676834106
step: 130, loss: 0.33132919669151306
step: 140, loss: 0.18099476397037506
step: 150, loss: 0.30576092004776
step: 160, loss: 0.24192854762077332
step: 170, loss: 0.3885268270969391
step: 180, loss: 0.1633080393075943
step: 190, loss: 0.15009328722953796
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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: 0.3650262653827667
step: 10, loss: 0.3041803240776062
step: 20, loss: 0.5739313960075378
step: 30, loss: 0.2317427694797516
step: 40, loss: 0.5480211973190308
step: 50, loss: 0.3148614466190338
step: 60, loss: 0.24676834046840668
step: 70, loss: 0.39511510729789734
step: 80, loss: 0.09524090588092804
step: 90, loss: 0.16212184727191925
step: 100, loss: 0.12736178934574127
step: 110, loss: 0.11598324775695801
step: 120, loss: 0.15423263609409332
step: 130, loss: 0.22185122966766357
step: 140, loss: 0.24751347303390503
step: 150, loss: 0.1451670378446579
step: 160, loss: 0.1380918174982071
step: 170, loss: 0.10088148713111877
step: 180, loss: 0.05261969938874245
step: 190, loss: 0.04880839213728905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7626666666666667, f1=0.7526881720430108,
best_f1=0.7526881720430108
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11898177862167358
step: 10, loss: 0.08005701005458832
step: 20, loss: 0.06007058173418045
step: 30, loss: 0.0048264795914292336
step: 40, loss: 0.1022893637418747
step: 50, loss: 0.1911298930644989

```

```

step: 60, loss: 0.048713576048612595
step: 70, loss: 0.24488240480422974
step: 80, loss: 0.18041932582855225
step: 90, loss: 0.09736145287752151
step: 100, loss: 0.07133448868989944
step: 110, loss: 0.41612300276756287
step: 120, loss: 0.1568625122308731
step: 130, loss: 0.12658676505088806
step: 140, loss: 0.044198449701070786
step: 150, loss: 0.16282323002815247
step: 160, loss: 0.06176392361521721
step: 170, loss: 0.37784963846206665
step: 180, loss: 0.13759993016719818
step: 190, loss: 0.07247236371040344
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8021680216802168, f1=0.7988980716253442,
best_f1=0.7988980716253442
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007094793953001499
step: 10, loss: 0.07404866814613342
step: 20, loss: 0.04646977782249451
step: 30, loss: 0.003707422409206629
step: 40, loss: 0.0720115676522255
step: 50, loss: 0.11362384259700775
step: 60, loss: 0.06282849609851837
step: 70, loss: 0.09135434031486511
step: 80, loss: 0.08626651763916016
step: 90, loss: 0.06688472628593445
step: 100, loss: 0.04965833202004433
step: 110, loss: 0.09303167462348938
step: 120, loss: 0.019341276958584785
step: 130, loss: 0.028603103011846542
step: 140, loss: 0.297396183013916
step: 150, loss: 0.1885388046503067
step: 160, loss: 0.010997519828379154
step: 170, loss: 0.03753634914755821
step: 180, loss: 0.03532928600907326
step: 190, loss: 0.012059759348630905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.827027027027027, f1=0.8181818181818181,
best_f1=0.8181818181818181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07032587379217148
step: 10, loss: 0.08266832679510117
step: 20, loss: 0.012602127157151699
step: 30, loss: 0.02771453745663166
step: 40, loss: 0.005952543579041958
step: 50, loss: 0.006925821769982576
step: 60, loss: 0.019587574526667595
step: 70, loss: 0.09585954248905182
step: 80, loss: 0.012423847801983356
step: 90, loss: 0.09004919230937958
step: 100, loss: 0.08258872479200363
step: 110, loss: 0.2389499396085739
step: 120, loss: 0.05855866149067879
step: 130, loss: 0.2211761474609375
step: 140, loss: 0.049462076276540756
step: 150, loss: 0.012870675884187222
step: 160, loss: 0.01331078726798296
step: 170, loss: 0.03140944242477417
step: 180, loss: 0.02093334309756756
step: 190, loss: 0.01668601855635643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.837465564738292, f1=0.8287292817679558,
best_f1=0.8287292817679558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18523406982421875
step: 10, loss: 0.03245275095105171
step: 20, loss: 0.014147532172501087
step: 30, loss: 0.00967487320303917
step: 40, loss: 0.008657197467982769
step: 50, loss: 0.02999255619943142

```



```

step: 60, loss: 0.07515013217926025
step: 70, loss: 0.05209222063422203
step: 80, loss: 0.02510799840092659
step: 90, loss: 0.004199173301458359
step: 100, loss: 0.0695725530385971
step: 110, loss: 0.010957148857414722
step: 120, loss: 0.02184871770441532
step: 130, loss: 0.08304891735315323
step: 140, loss: 0.0023014643229544163
step: 150, loss: 0.001411794451996684
step: 160, loss: 0.08521738648414612
step: 170, loss: 0.1610368937253952
step: 180, loss: 0.009984476491808891
step: 190, loss: 0.010709665715694427
/usr/local/lib/python3.7/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=0.8306878306878307, f1=0.8152173913043478,
best_f1=0.8287292817679558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017023224383592606
step: 10, loss: 0.004047900438308716
step: 20, loss: 0.0005991618381813169
step: 30, loss: 0.011638902127742767
step: 40, loss: 0.0025495088193565607
step: 50, loss: 0.002631198614835739
step: 60, loss: 0.03366321325302124
step: 70, loss: 0.0010537935886532068
step: 80, loss: 0.0009469452779740095
step: 90, loss: 0.006812055129557848
step: 100, loss: 0.14576603472232819
step: 110, loss: 0.004054867196828127
step: 120, loss: 0.10058264434337616
step: 130, loss: 0.013733171857893467
step: 140, loss: 0.06082477420568466
step: 150, loss: 0.002522974042221904
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.11701317131519318
step: 170, loss: 0.037385616451501846
step: 180, loss: 0.008597473613917828
step: 190, loss: 0.008424720726907253
/usr/local/lib/python3.7/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=0.8196721311475409, f1=0.8314606741573033,
best_f1=0.8287292817679558

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04463720694184303
step: 10, loss: 0.002842289162799716
step: 20, loss: 0.0007523676031269133
step: 30, loss: 0.10123808681964874
step: 40, loss: 0.12212439626455307
step: 50, loss: 0.11077845096588135
step: 60, loss: 0.04548794403672218
step: 70, loss: 0.0007730043726041913
step: 80, loss: 0.07510115951299667
step: 90, loss: 0.0213436558842659
step: 100, loss: 0.12064991146326065
step: 110, loss: 0.00449054641649127
step: 120, loss: 0.004073929041624069
step: 130, loss: 0.006664691027253866
step: 140, loss: 0.0016501586651429534
step: 150, loss: 0.06217131018638611
step: 160, loss: 0.00980918575078249
step: 170, loss: 0.0011160534340888262
step: 180, loss: 0.07993227988481522
step: 190, loss: 0.05768207460641861
/usr/local/lib/python3.7/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=0.8292682926829268, f1=0.8301886792452831,
best_f1=0.8287292817679558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007725170347839594
step: 10, loss: 0.1088959127664566
step: 20, loss: 0.04719943180680275
step: 30, loss: 0.004607769660651684
step: 40, loss: 0.028774669393897057
step: 50, loss: 0.003225291846320033
step: 60, loss: 0.012311390601098537
step: 70, loss: 0.013909085653722286
step: 80, loss: 0.0022890325635671616
step: 90, loss: 0.004130674060434103
step: 100, loss: 0.0037552486173808575
step: 110, loss: 0.004182795528322458
step: 120, loss: 0.03681932017207146
step: 130, loss: 0.0036776778288185596
step: 140, loss: 0.04683981090784073
step: 150, loss: 0.0031600745860487223
step: 160, loss: 0.2523016929626465

```

```

step: 170, loss: 0.151840478181839
step: 180, loss: 0.07094531506299973
step: 190, loss: 0.001279487507417798
/usr/local/lib/python3.7/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=0.8346883468834688, f1=0.8310249307479225,
best_f1=0.8287292817679558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033157109282910824
step: 10, loss: 0.003599768038839102
step: 20, loss: 0.005637108813971281
step: 30, loss: 0.06741087883710861
step: 40, loss: 0.027494044974446297
step: 50, loss: 0.00881132297217846
step: 60, loss: 0.0035740884486585855
step: 70, loss: 0.0012668668059632182
step: 80, loss: 0.014960313215851784
step: 90, loss: 0.001688977936282754
step: 100, loss: 0.0007111984305083752
step: 110, loss: 0.0019391984678804874
step: 120, loss: 0.0007886101375333965
step: 130, loss: 0.2463146299123764
step: 140, loss: 0.016789915040135384
step: 150, loss: 0.0063246507197618484
step: 160, loss: 0.001174303935840726
step: 170, loss: 0.0034805091563612223
step: 180, loss: 0.004181490745395422
step: 190, loss: 0.0020016611088067293
/usr/local/lib/python3.7/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=0.8241469816272966, f1=0.8494623655913978,
best_f1=0.8287292817679558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045379724353551865
step: 10, loss: 0.0003839465498458594
step: 20, loss: 0.000271008990239352
step: 30, loss: 0.00654191616922617
step: 40, loss: 0.0006513044936582446
step: 50, loss: 0.0026281706523150206
step: 60, loss: 0.007511754985898733
step: 70, loss: 0.0010722840670496225
step: 80, loss: 0.02655818872153759

```

```

step: 90, loss: 0.0005359596107155085
step: 100, loss: 0.010926254093647003
step: 110, loss: 0.12403841316699982
step: 120, loss: 0.0016905952943488955
step: 130, loss: 0.0069246371276676655
step: 140, loss: 0.000972090580035001
step: 150, loss: 0.0012868117773905396
step: 160, loss: 0.0009168954566121101
step: 170, loss: 0.0006811570492573082
step: 180, loss: 0.0006586131057702005
step: 190, loss: 0.0005335789755918086
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8378378378378378, 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: 0.03451388701796532
step: 10, loss: 0.00887366198003292
step: 20, loss: 0.0015828789910301566
step: 30, loss: 0.005957919172942638
step: 40, loss: 0.0002954691299237311
step: 50, loss: 0.0008890433236956596
step: 60, loss: 0.0006819875561632216
step: 70, loss: 0.001789251109585166
step: 80, loss: 0.01573259010910988
step: 90, loss: 0.0005747483810409904
step: 100, loss: 0.0021692991722375154
step: 110, loss: 0.024741077795624733
step: 120, loss: 0.007784458342939615
step: 130, loss: 0.0026649460196495056
step: 140, loss: 0.03445683792233467
step: 150, loss: 0.002988675143569708
step: 160, loss: 0.0005011627799831331
step: 170, loss: 0.012152928858995438
step: 180, loss: 0.0005216379649937153
step: 190, loss: 0.028602221980690956
/usr/local/lib/python3.7/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=0.8378378378378378, f1=0.8388888888888889,
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: 0.0027191233821213245
step: 10, loss: 0.005317302420735359
step: 20, loss: 0.00024434737861156464
step: 30, loss: 0.008236239664256573
step: 40, loss: 0.002934823976829648
step: 50, loss: 0.0026698592118918896
step: 60, loss: 0.00023351729032583535
step: 70, loss: 0.024143682792782784
step: 80, loss: 0.0007603802951052785
step: 90, loss: 0.013447328470647335
step: 100, loss: 0.0001853640569606796
step: 110, loss: 0.001045578857883811
step: 120, loss: 0.008218144997954369
step: 130, loss: 0.0007019509212113917
step: 140, loss: 0.0015009205089882016
step: 150, loss: 0.00019461323972791433
step: 160, loss: 0.0004791322280652821
step: 170, loss: 0.0049142721109092236
step: 180, loss: 0.0013296952238306403
step: 190, loss: 0.15780290961265564

/usr/local/lib/python3.7/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=0.8337874659400545, f1=0.8365650969529086,
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: 0.0008888756274245679
step: 10, loss: 0.01157804299145937
step: 20, loss: 0.0008261362672783434
step: 30, loss: 0.0023805666714906693
step: 40, loss: 0.001639176276512444
step: 50, loss: 0.00450216606259346
step: 60, loss: 0.00073719717329368
step: 70, loss: 0.00552136218175292
step: 80, loss: 0.0010095834732055664
step: 90, loss: 0.00044572007027454674
step: 100, loss: 0.0005328685510903597
step: 110, loss: 0.0002614022814668715
step: 120, loss: 0.0018307516584172845
step: 130, loss: 0.0005268256063573062
step: 140, loss: 0.018439197912812233
step: 150, loss: 0.07658643275499344
step: 160, loss: 0.00038577785016968846

```

step: 170, loss: 0.0002466402074787766
step: 180, loss: 0.002795073203742504
step: 190, loss: 0.0011604016181081533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8387096774193548, f1=0.8455284552845528,
best_f1=0.8455284552845528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002316474710823968
step: 10, loss: 0.16408097743988037
step: 20, loss: 0.006955675780773163
step: 30, loss: 0.0008304372895509005
step: 40, loss: 0.0005875880597159266
step: 50, loss: 0.00020437882631085813
step: 60, loss: 0.07977007329463959
step: 70, loss: 0.0027736180927604437
step: 80, loss: 0.08880136162042618
step: 90, loss: 0.0007292331429198384
step: 100, loss: 0.0015169387916103005
step: 110, loss: 0.0005731989513151348
step: 120, loss: 0.0032560951076447964
step: 130, loss: 0.0005771215073764324
step: 140, loss: 0.00377363758161664
step: 150, loss: 0.001977561041712761
step: 160, loss: 0.0008159191347658634
step: 170, loss: 0.001860042568296194
step: 180, loss: 0.0003352165222167969
step: 190, loss: 0.0012737351935356855
/usr/local/lib/python3.7/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=0.8359788359788362, f1=0.8525469168900803,
best_f1=0.8455284552845528
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 [00:12, 157.78it/s]
load_f1 = 0.7999999999999999
real_f1 = 0.7936507936507936
733it [00:00, 3165.31it/s]/usr/local/lib/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, 134.51it/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: 0.4804520606994629
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.42273518443107605
step: 20, loss: 0.2990303635597229
step: 30, loss: 0.3982568681240082
step: 40, loss: 0.5750716924667358
step: 50, loss: 0.3462660610675812
step: 60, loss: 0.5626756548881531
step: 70, loss: 0.31200194358825684
step: 80, loss: 0.2355489879846573
step: 90, loss: 0.23390638828277588
step: 100, loss: 0.14730164408683777
step: 110, loss: 0.37903892993927
step: 120, loss: 0.3294135630130768
step: 130, loss: 0.30141404271125793
step: 140, loss: 0.41021308302879333
step: 150, loss: 0.29631438851356506
step: 160, loss: 0.36705267429351807
step: 170, loss: 0.31875795125961304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.19479905437352246, f1=0.19458946369245375,
best_f1=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32747143507003784
step: 10, loss: 0.49588844180107117
step: 20, loss: 0.33177316188812256
step: 30, loss: 0.3237922787666321
step: 40, loss: 0.09756912291049957
step: 50, loss: 0.47199147939682007

```

```

step: 60, loss: 0.1975613385438919
step: 70, loss: 0.49816539883613586
step: 80, loss: 0.2381219118833542
step: 90, loss: 0.23430098593235016
step: 100, loss: 0.5336788892745972
step: 110, loss: 0.2647239565849304
step: 120, loss: 0.2545692026615143
step: 130, loss: 0.552038311958313
step: 140, loss: 0.5509451031684875
step: 150, loss: 0.4448794722557068
step: 160, loss: 0.4508115351200104
step: 170, loss: 0.3911588788032532
/usr/local/lib/python3.7/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=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6082710027694702
step: 10, loss: 0.304830938577652
step: 20, loss: 0.23096835613250732
step: 30, loss: 0.2358989119529724
step: 40, loss: 0.369495689868927
step: 50, loss: 0.5645074248313904
step: 60, loss: 0.30729931592941284
step: 70, loss: 0.24257366359233856
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.3734528422355652
step: 90, loss: 0.5164515376091003
step: 100, loss: 0.3174067437648773
step: 110, loss: 0.1799658238887787
step: 120, loss: 0.5594722628593445
step: 130, loss: 0.5661367177963257
step: 140, loss: 0.4366133213043213
step: 150, loss: 0.20167888700962067
step: 160, loss: 0.16159141063690186
step: 170, loss: 0.309758722782135
/usr/local/lib/python3.7/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=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40537166595458984

```

```

step: 10, loss: 0.5287166833877563
step: 20, loss: 0.24718952178955078
step: 30, loss: 0.4469950497150421
step: 40, loss: 0.2559632956981659
step: 50, loss: 0.385043740272522
step: 60, loss: 0.7208488583564758
step: 70, loss: 0.3091242015361786
step: 80, loss: 0.4506172835826874
step: 90, loss: 0.30657973885536194
step: 100, loss: 0.36794155836105347
step: 110, loss: 0.44356968998908997
step: 120, loss: 0.47196534276008606
step: 130, loss: 0.32307466864585876
step: 140, loss: 0.23826487362384796
step: 150, loss: 0.7098245620727539
step: 160, loss: 0.11688109487295151
step: 170, loss: 0.24037186801433563
/usr/local/lib/python3.7/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=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3838367164134979
step: 10, loss: 0.30612146854400635
step: 20, loss: 0.3177052140235901
step: 30, loss: 0.3361036479473114
step: 40, loss: 0.25107914209365845
step: 50, loss: 0.2356126457452774
step: 60, loss: 0.31132447719573975
step: 70, loss: 0.3141709566116333
step: 80, loss: 0.07454885542392731
step: 90, loss: 0.6290559768676758
step: 100, loss: 0.256600558757782
step: 110, loss: 0.26451560854911804
step: 120, loss: 0.13183245062828064
step: 130, loss: 0.24690082669258118
step: 140, loss: 0.1740707904100418
step: 150, loss: 0.30771660804748535
step: 160, loss: 0.23816415667533875
step: 170, loss: 0.3741769790649414
/usr/local/lib/python3.7/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=0.19458946369245375

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31771084666252136
step: 10, loss: 0.11110813915729523
step: 20, loss: 0.39426422119140625
step: 30, loss: 0.25276103615760803
step: 40, loss: 0.43546685576438904
step: 50, loss: 0.27123257517814636
step: 60, loss: 0.37838199734687805
step: 70, loss: 0.3728693425655365
step: 80, loss: 0.19128161668777466
step: 90, loss: 0.33277323842048645
step: 100, loss: 0.19203388690948486
step: 110, loss: 0.44259944558143616
step: 120, loss: 0.4404819905757904
step: 130, loss: 0.4811154305934906
step: 140, loss: 0.2592393159866333
step: 150, loss: 0.16978827118873596
step: 160, loss: 0.3141945004463196
step: 170, loss: 0.24036332964897156
/usr/local/lib/python3.7/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=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23891912400722504
step: 10, loss: 0.4637157917022705
step: 20, loss: 0.44777464866638184
step: 30, loss: 0.51718670129776
step: 40, loss: 0.12282392382621765
step: 50, loss: 0.4447418749332428
step: 60, loss: 0.5495181679725647
step: 70, loss: 0.3808313012123108
step: 80, loss: 0.1848193258047104
step: 90, loss: 0.4766998589038849
step: 100, loss: 0.31953710317611694
step: 110, loss: 0.3801144063472748
step: 120, loss: 0.38490262627601624
step: 130, loss: 0.1957695633172989
step: 140, loss: 0.12531529366970062
step: 150, loss: 0.30618342757225037
step: 160, loss: 0.3865160346031189
step: 170, loss: 0.3148181438446045
/usr/local/lib/python3.7/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=0.19458946369245375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44865262508392334
step: 10, loss: 0.5131452083587646
step: 20, loss: 0.24516384303569794
step: 30, loss: 0.3175791800022125
step: 40, loss: 0.24695433676242828
step: 50, loss: 0.3802810311317444
step: 60, loss: 0.3865628242492676
step: 70, loss: 0.1861758530139923
step: 80, loss: 0.31042054295539856
step: 90, loss: 0.5782753825187683
step: 100, loss: 0.25020110607147217
step: 110, loss: 0.5204786062240601
step: 120, loss: 0.3106538653373718
step: 130, loss: 0.31548115611076355
step: 140, loss: 0.5068181753158569
step: 150, loss: 0.24116811156272888
step: 160, loss: 0.17658349871635437
step: 170, loss: 0.4510727822780609
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20526315789473684, f1=0.20995079278294151,
best_f1=0.20995079278294151
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37710443139076233
step: 10, loss: 0.6023570895195007
step: 20, loss: 0.37756648659706116
step: 30, loss: 0.17330877482891083
step: 40, loss: 0.37281322479248047
step: 50, loss: 0.24669025838375092
step: 60, loss: 0.3815438151359558
step: 70, loss: 0.33414262533187866
step: 80, loss: 0.1605762392282486
step: 90, loss: 0.4444359838962555
step: 100, loss: 0.4977905750274658

```

```

step: 110, loss: 0.30602261424064636
step: 120, loss: 0.4330407381057739
step: 130, loss: 0.3166835308074951
step: 140, loss: 0.3746078610420227
step: 150, loss: 0.3610338568687439
step: 160, loss: 0.25164613127708435
step: 170, loss: 0.42729005217552185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.30180180180180183, f1=0.32125984251968503,
best_f1=0.32125984251968503
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.32880115509033203
step: 10, loss: 0.23199214041233063
step: 20, loss: 0.26159411668777466
step: 30, loss: 0.6676998734474182
step: 40, loss: 0.2971704304218292
step: 50, loss: 0.4168948531150818
step: 60, loss: 0.4122065007686615
step: 70, loss: 0.29941892623901367
step: 80, loss: 0.21754330396652222
step: 90, loss: 0.2175407111644745
step: 100, loss: 0.11927659809589386
step: 110, loss: 0.15841686725616455
step: 120, loss: 0.2149120271205902
step: 130, loss: 0.10608713328838348
step: 140, loss: 0.09036378562450409
step: 150, loss: 0.18493962287902832
step: 160, loss: 0.09637416899204254
step: 170, loss: 0.32733017206192017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7201946472019466, f1=0.7446300715990454,
best_f1=0.7446300715990454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42520222067832947
step: 10, loss: 0.08208047598600388
step: 20, loss: 0.1931377798318863
step: 30, loss: 0.3314301669597626
step: 40, loss: 0.05613846704363823
step: 50, loss: 0.08341236412525177
step: 60, loss: 0.35123246908187866
step: 70, loss: 0.07454244792461395
step: 80, loss: 0.041757553815841675
step: 90, loss: 0.033574532717466354
step: 100, loss: 0.5633094906806946
step: 110, loss: 0.2871103584766388
step: 120, loss: 0.09681368619203568
step: 130, loss: 0.13852418959140778
step: 140, loss: 0.161291241645813
step: 150, loss: 0.06401710957288742
step: 160, loss: 0.18438871204853058
step: 170, loss: 0.1642344743013382
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7494145199063231, f1=0.751131221719457,
best_f1=0.751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035216353833675385
step: 10, loss: 0.0959884449839592
step: 20, loss: 0.053299471735954285
step: 30, loss: 0.26075223088264465
step: 40, loss: 0.04527517035603523
step: 50, loss: 0.12202531844377518
step: 60, loss: 0.17705464363098145
step: 70, loss: 0.10333328694105148
step: 80, loss: 0.027962304651737213
step: 90, loss: 0.20476914942264557
step: 100, loss: 0.13330091536045074
step: 110, loss: 0.11889555305242538
step: 120, loss: 0.09725453704595566
step: 130, loss: 0.28505679965019226
step: 140, loss: 0.06754545867443085
step: 150, loss: 0.07568835467100143
step: 160, loss: 0.30598002672195435

```

```

step: 170, loss: 0.22888995707035065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7961165048543688, f1=0.8175519630484988,
best_f1=0.8175519630484988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11634838581085205
step: 10, loss: 0.014463079161942005
step: 20, loss: 0.06937643885612488
step: 30, loss: 0.09560857713222504
step: 40, loss: 0.130165696144104
step: 50, loss: 0.09455594420433044
step: 60, loss: 0.13005399703979492
step: 70, loss: 0.28889763355255127
step: 80, loss: 0.04110514372587204
step: 90, loss: 0.04207786172628403
step: 100, loss: 0.03889577090740204
step: 110, loss: 0.041881900280714035
step: 120, loss: 0.1595037430524826
step: 130, loss: 0.056333377957344055
step: 140, loss: 0.15878750383853912
step: 150, loss: 0.10559035837650299
step: 160, loss: 0.05834846943616867
step: 170, loss: 0.010576016269624233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8179551122194514, f1=0.8439024390243902,
best_f1=0.8439024390243902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04281402751803398
step: 10, loss: 0.029486997053027153
step: 20, loss: 0.06642124801874161
step: 30, loss: 0.04061717167496681
step: 40, loss: 0.1251949965953827

```



```

step: 50, loss: 0.13892634212970734
step: 60, loss: 0.11117345839738846
step: 70, loss: 0.06437815725803375
step: 80, loss: 0.04890303313732147
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.027987133711576462
step: 100, loss: 0.2340114265680313
step: 110, loss: 0.14640356600284576
step: 120, loss: 0.1837691217660904
step: 130, loss: 0.14623215794563293
step: 140, loss: 0.10070439428091049
step: 150, loss: 0.03868776187300682
step: 160, loss: 0.0672406256198883
step: 170, loss: 0.08596747368574142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8214285714285714, f1=0.8486352357320099,
best_f1=0.8486352357320099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005744646769016981
step: 10, loss: 0.030673211440443993
step: 20, loss: 0.12226235121488571
step: 30, loss: 0.07392329722642899
step: 40, loss: 0.0440593883395195
step: 50, loss: 0.012800746597349644
step: 60, loss: 0.020810144022107124
step: 70, loss: 0.1193031370639801
step: 80, loss: 0.03374774381518364
step: 90, loss: 0.019090969115495682
step: 100, loss: 0.14311400055885315
step: 110, loss: 0.25711488723754883
step: 120, loss: 0.043224699795246124
step: 130, loss: 0.10583695024251938
step: 140, loss: 0.03819388896226883
step: 150, loss: 0.008944619446992874
step: 160, loss: 0.0757087841629982
step: 170, loss: 0.016163701191544533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.8256410256410256, f1=0.8493827160493828,
best_f1=0.8493827160493828
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:08, 226.04it/s]
load_f1 = 0.5153583617747441
real_f1 = 0.4606580829756796
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, 137.10it/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, 408kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 802kB/s]
Downloading: 100% 456k/456k [00:00<00:00, 513kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 69.1MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

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

```

opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5603840947151184
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: 0.41869428753852844
step: 20, loss: 0.483099102973938
step: 30, loss: 0.3204706609249115
step: 40, loss: 0.3254481256008148
step: 50, loss: 0.5180754065513611
step: 60, loss: 0.24372941255569458
step: 70, loss: 0.09392748773097992
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.20864108204841614
step: 90, loss: 0.18466155230998993
step: 100, loss: 0.19864369928836823
step: 110, loss: 0.09101590514183044
step: 120, loss: 0.05758127197623253
step: 130, loss: 0.07573363184928894
step: 140, loss: 0.08038432151079178
step: 150, loss: 0.1890588104724884
step: 160, loss: 0.07375926524400711
step: 170, loss: 0.19376909732818604
step: 180, loss: 0.11958868056535721
step: 190, loss: 0.02244962379336357
step: 200, loss: 0.04007464647293091
step: 210, loss: 0.018577812239527702
step: 220, loss: 0.06763537228107452
step: 230, loss: 0.03389210253953934

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9522752497225305, f1=0.9652855543113102,
best_f1=0.9652855543113102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005216829478740692
step: 10, loss: 0.16116321086883545
step: 20, loss: 0.007692575454711914
step: 30, loss: 0.027728864923119545
step: 40, loss: 0.06074342131614685
step: 50, loss: 0.02735028602182865
step: 60, loss: 0.00962441973388195
step: 70, loss: 0.012331780046224594
step: 80, loss: 0.014047008939087391
step: 90, loss: 0.0537702701985836
step: 100, loss: 0.21612638235092163
step: 110, loss: 0.04245621711015701
step: 120, loss: 0.010030413046479225
step: 130, loss: 0.062412627041339874
step: 140, loss: 0.004133778624236584
step: 150, loss: 0.11271022260189056
step: 160, loss: 0.013297543860971928
step: 170, loss: 0.005257719662040472
step: 180, loss: 0.012711008079349995
step: 190, loss: 0.006774731446057558
step: 200, loss: 0.024369770660996437
step: 210, loss: 0.015391724184155464
step: 220, loss: 0.010499070398509502
step: 230, loss: 0.005288266111165285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9617977528089887, f1=0.9661399548532732,
best_f1=0.9661399548532732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.008234823122620583
step: 10, loss: 0.08130664378404617
step: 20, loss: 0.011239803396165371
step: 30, loss: 0.005936162546277046
step: 40, loss: 0.12243099510669708
step: 50, loss: 0.018985947594046593
step: 60, loss: 0.03142521530389786
step: 70, loss: 0.003878724528476596
step: 80, loss: 0.03484528139233589
step: 90, loss: 0.0073627387173473835
step: 100, loss: 0.016836419701576233
step: 110, loss: 0.02415693923830986
step: 120, loss: 0.0004767464997712523
step: 130, loss: 0.032540418207645416
step: 140, loss: 0.009684291668236256
step: 150, loss: 0.03702286258339882
step: 160, loss: 0.0023036908823996782
step: 170, loss: 0.000826281844638288
step: 180, loss: 0.03992520272731781
step: 190, loss: 0.022837771102786064
step: 200, loss: 0.018007785081863403
step: 210, loss: 0.0007097060442902148
step: 220, loss: 0.013705874793231487
step: 230, loss: 0.010858849622309208
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.971815107102593, f1=0.9717514124293786,
best_f1=0.9717514124293786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00866026896983385
step: 10, loss: 0.005324407946318388
step: 20, loss: 0.011654647998511791
step: 30, loss: 0.004792242310941219
step: 40, loss: 0.005124376155436039
step: 50, loss: 0.004621647298336029
step: 60, loss: 0.00875469483435154
step: 70, loss: 0.012346002273261547
step: 80, loss: 0.026886170729994774
step: 90, loss: 0.02361447550356388
step: 100, loss: 0.007880236953496933
step: 110, loss: 0.02787969261407852

```

```

step: 120, loss: 0.04064633697271347
step: 130, loss: 0.02557808719575405
step: 140, loss: 0.05397799611091614
step: 150, loss: 0.006978299934417009
step: 160, loss: 0.0017358734039589763
step: 170, loss: 0.002293243305757642
step: 180, loss: 0.13616317510604858
step: 190, loss: 0.003246505279093981
step: 200, loss: 0.02833745814859867
step: 210, loss: 0.12921002507209778
step: 220, loss: 0.002028920454904437
step: 230, loss: 0.01415198016911745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.972129319955407, f1=0.9765363128491621,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00503005925565958
step: 10, loss: 0.0032687231432646513
step: 20, loss: 0.0792372077703476
step: 30, loss: 0.004978867247700691
step: 40, loss: 0.0011768965050578117
step: 50, loss: 0.0018860847922042012
step: 60, loss: 0.03416096419095993
step: 70, loss: 0.006297217216342688
step: 80, loss: 0.05900948494672775
step: 90, loss: 0.05796784162521362
step: 100, loss: 0.00037532500573433936
step: 110, loss: 0.01710425317287445
step: 120, loss: 0.002459881594404578
step: 130, loss: 0.002251954050734639
step: 140, loss: 0.008327201008796692
step: 150, loss: 0.016157226637005806
step: 160, loss: 0.001754065859131515
step: 170, loss: 0.002534942002967
step: 180, loss: 0.0029887931887060404
step: 190, loss: 0.01905428059399128
step: 200, loss: 0.005317640025168657
step: 210, loss: 0.0026821347419172525
step: 220, loss: 0.007493558805435896
step: 230, loss: 0.007478382904082537

```

```

/usr/local/lib/python3.7/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=0.9689578713968958, f1=0.9720670391061451,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017763498472049832
step: 10, loss: 0.01559522282332182
step: 20, loss: 0.0014207131462171674
step: 30, loss: 0.001432748045772314
step: 40, loss: 0.00127495510969311
step: 50, loss: 0.0008163530728779733
step: 60, loss: 0.0016580767696723342
step: 70, loss: 0.09705974161624908
step: 80, loss: 0.0009796046651899815
step: 90, loss: 0.002808136399835348
step: 100, loss: 0.001010407111607492
step: 110, loss: 0.08410480618476868
step: 120, loss: 0.0005185177433304489
step: 130, loss: 0.0022665783762931824
step: 140, loss: 0.005692286416888237
step: 150, loss: 0.0002956808020826429
step: 160, loss: 0.03944162279367447
step: 170, loss: 0.005426329560577869
step: 180, loss: 0.01175080705434084
step: 190, loss: 0.001438019098713994
step: 200, loss: 0.04337392747402191
step: 210, loss: 0.005727787967771292
step: 220, loss: 0.0100092189386487
step: 230, loss: 0.0005460705724544823
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753914988814317, f1=0.9787709497206705,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012880993308499455
step: 10, loss: 0.00032674852991476655
step: 20, loss: 0.0002631522947922349
step: 30, loss: 0.00018079840810969472

```



```

step: 40, loss: 0.0046778530813753605
step: 50, loss: 0.00027910067001357675
step: 60, loss: 0.003481767838820815
step: 70, loss: 0.0007012051064521074
step: 80, loss: 0.00034515850711613894
step: 90, loss: 0.0021177460439503193
step: 100, loss: 0.0005404918338172138
step: 110, loss: 0.0005990279605612159
step: 120, loss: 0.0028245174326002598
step: 130, loss: 0.004684338811784983
step: 140, loss: 0.0003790563205257058
step: 150, loss: 0.03668540343642235
step: 160, loss: 0.0030688666738569736
step: 170, loss: 0.013353785499930382
step: 180, loss: 0.0030765277333557606
step: 190, loss: 0.0060286205261945724
step: 200, loss: 0.0012516287388280034
step: 210, loss: 0.015184126794338226
step: 220, loss: 0.0002514387888368219
step: 230, loss: 0.0006395963137038052
/usr/local/lib/python3.7/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=0.9753363228699552, f1=0.9766407119021134,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006360829574987292
step: 10, loss: 0.004830750171095133
step: 20, loss: 0.000811594189144671
step: 30, loss: 0.0013510023709386587
step: 40, loss: 0.03188563510775566
step: 50, loss: 0.005887380335479975
step: 60, loss: 0.007982463575899601
step: 70, loss: 0.00028020376339554787
step: 80, loss: 0.057586442679166794
step: 90, loss: 0.0001711068325676024
step: 100, loss: 0.000508356315549463
step: 110, loss: 0.00434210617095232
step: 120, loss: 0.0010683861328288913
step: 130, loss: 0.008329296484589577
step: 140, loss: 0.002192300045862794
step: 150, loss: 0.16006219387054443
step: 160, loss: 0.00045706095988862216
step: 170, loss: 0.005799020174890757
step: 180, loss: 0.0011925242142751813
step: 190, loss: 0.001286056125536561

```

```

step: 200, loss: 0.02060036174952984
step: 210, loss: 0.005709701683372259
step: 220, loss: 0.00025593661121092737
step: 230, loss: 0.0008402615203522146
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9799107142857142,
best_f1=0.9799107142857142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001152471304521896
step: 10, loss: 0.0004819343739654869
step: 20, loss: 0.00042839342495426536
step: 30, loss: 0.0003324792196508497
step: 40, loss: 0.0002025310677709058
step: 50, loss: 0.0005612607346847653
step: 60, loss: 0.0003590538108255714
step: 70, loss: 0.027364538982510567
step: 80, loss: 0.00023935105127748102
step: 90, loss: 0.06827179342508316
step: 100, loss: 0.0011531459167599678
step: 110, loss: 0.0007130114245228469
step: 120, loss: 0.025854293256998062
step: 130, loss: 0.0022754953242838383
step: 140, loss: 0.001808198052458465
step: 150, loss: 0.00018890078354161233
step: 160, loss: 0.021969741210341454
step: 170, loss: 0.016134539619088173
step: 180, loss: 0.00020372839935589582
step: 190, loss: 5.714933286071755e-05
step: 200, loss: 9.867778135230765e-05
step: 210, loss: 0.05071471631526947
step: 220, loss: 0.00029358192114159465
step: 230, loss: 0.007784538436681032
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798206278026906, f1=0.976324689966178,

```

```

best_f1=0.976324689966178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002264420036226511
step: 10, loss: 0.0008713314309716225
step: 20, loss: 0.0009231443400494754
step: 30, loss: 0.00047418917529284954
step: 40, loss: 0.0005329434061422944
step: 50, loss: 0.0006559651810675859
step: 60, loss: 0.00012765510473400354
step: 70, loss: 0.0006947565707378089
step: 80, loss: 5.8947345678461716e-05
step: 90, loss: 0.0005800375947728753
step: 100, loss: 0.000186852557817474
step: 110, loss: 0.011939306743443012
step: 120, loss: 4.667346365749836e-05
step: 130, loss: 0.0035627014003694057
step: 140, loss: 7.509940041927621e-05
step: 150, loss: 0.011209327727556229
step: 160, loss: 0.0003201059007551521
step: 170, loss: 0.00011877997167175636
step: 180, loss: 0.0003911773092113435
step: 190, loss: 0.0049023558385670185
step: 200, loss: 0.00021036452380940318
step: 210, loss: 0.0030092992819845676
step: 220, loss: 0.001648055505938828
step: 230, loss: 0.0001537299540359527
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9830124575311437, 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: 7.011082925600931e-05
step: 10, loss: 8.847279968904331e-05
step: 20, loss: 8.88678987394087e-05
step: 30, loss: 0.00024034381203819066
step: 40, loss: 4.6261120587587357e-05
step: 50, loss: 0.00011568655463634059
step: 60, loss: 0.00753629207611084
step: 70, loss: 0.0006896804552525282

```

```

step: 80, loss: 0.0003360798582434654
step: 90, loss: 0.04360875114798546
step: 100, loss: 0.00012354592035990208
step: 110, loss: 0.00015708457794971764
step: 120, loss: 0.0001341735478490591
step: 130, loss: 0.0005835068877786398
step: 140, loss: 0.008711481466889381
step: 150, loss: 0.0012147376546636224
step: 160, loss: 0.010426956228911877
step: 170, loss: 0.0013775593833997846
step: 180, loss: 0.0005522966966964304
step: 190, loss: 9.733975457493216e-05
step: 200, loss: 0.006571334786713123
step: 210, loss: 9.780841355677694e-05
step: 220, loss: 0.0006095438147895038
step: 230, loss: 0.0002332760050194338
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9753363228699552,
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: 0.00014177514822222292
step: 10, loss: 9.392580977873877e-05
step: 20, loss: 0.029379896819591522
step: 30, loss: 0.03457018733024597
step: 40, loss: 0.0001937846391228959
step: 50, loss: 0.0007960163638927042
step: 60, loss: 0.00018259880016557872
step: 70, loss: 7.08842562744394e-05
step: 80, loss: 0.0001128051953855902
step: 90, loss: 0.00033044294104911387
step: 100, loss: 4.582671317621134e-05
step: 110, loss: 5.956737004453316e-05
step: 120, loss: 0.0062574963085353374
step: 130, loss: 0.00010990120790665969
step: 140, loss: 0.00013762518938165158
step: 150, loss: 6.937637954251841e-05
step: 160, loss: 0.0006135979201644659
step: 170, loss: 0.003622936550527811
step: 180, loss: 5.763283479609527e-05
step: 190, loss: 0.0001912233274197206
step: 200, loss: 6.382275751093403e-05
step: 210, loss: 0.0033558157738298178
step: 220, loss: 0.0013293945230543613
step: 230, loss: 9.998594759963453e-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=0.9809203142536477, f1=0.9741282339707535,
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: 0.0017859541112557054
step: 10, loss: 8.645420894026756e-05
step: 20, loss: 0.001471543568186462
step: 30, loss: 0.0003341937263030559
step: 40, loss: 0.000699238502420485
step: 50, loss: 0.0036049021873623133
step: 60, loss: 8.282971248263493e-05
step: 70, loss: 0.0001294098183279857
step: 80, loss: 0.0003443494497332722
step: 90, loss: 6.914211553521454e-05
step: 100, loss: 7.510363502660766e-05
step: 110, loss: 0.0003282371908426285
step: 120, loss: 0.0004848005482926965
step: 130, loss: 5.673907435266301e-05
step: 140, loss: 0.0001403245551045984
step: 150, loss: 6.801645940868184e-05
step: 160, loss: 0.01010978128761053
step: 170, loss: 4.636175435734913e-05
step: 180, loss: 0.02511814422905445
step: 190, loss: 5.2791037887800485e-05
step: 200, loss: 1.9952338334405795e-05
step: 210, loss: 0.0002337643672944978
step: 220, loss: 3.811530041275546e-05
step: 230, loss: 0.0010031303390860558
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9787234042553192,
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: 3.058358561247587e-05
step: 10, loss: 3.6185905628371984e-05
step: 20, loss: 0.00017296394798904657
step: 30, loss: 0.0005121623398736119
step: 40, loss: 0.00015772617189213634
step: 50, loss: 4.476205867831595e-05
step: 60, loss: 3.2658525014994666e-05
step: 70, loss: 4.8183206672547385e-05

```

```

step: 80, loss: 4.374898344394751e-05
step: 90, loss: 3.969088720623404e-05
step: 100, loss: 3.358594767632894e-05
step: 110, loss: 0.00024874668451957405
step: 120, loss: 2.7730278816306964e-05
step: 130, loss: 0.0003109117387793958
step: 140, loss: 0.00022890089894644916
step: 150, loss: 3.174886660417542e-05
step: 160, loss: 3.7180347135290504e-05
step: 170, loss: 0.0003215441247448325
step: 180, loss: 2.8378555725794286e-05
step: 190, loss: 3.20513172482606e-05
step: 200, loss: 4.4094074837630615e-05
step: 210, loss: 2.998375020979438e-05
step: 220, loss: 8.576698019169271e-05
step: 230, loss: 0.0001401438203174621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9810901001112348, f1=0.9766925638179801,
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: 0.013320636004209518
step: 10, loss: 4.188122693449259e-05
step: 20, loss: 0.00011844252003356814
step: 30, loss: 4.179216193733737e-05
step: 40, loss: 3.3153544791275635e-05
step: 50, loss: 6.511755054816604e-05
step: 60, loss: 0.028338691219687462
step: 70, loss: 0.0007436834857799113
step: 80, loss: 0.02213863655924797
step: 90, loss: 3.137662497465499e-05
step: 100, loss: 2.634464362927247e-05
step: 110, loss: 4.465689562493935e-05
step: 120, loss: 0.022374557331204414
step: 130, loss: 3.57240169250872e-05
step: 140, loss: 0.0037790792994201183
step: 150, loss: 0.00014220993034541607
step: 160, loss: 0.011501915752887726
step: 170, loss: 2.7164081984665245e-05
step: 180, loss: 4.1795417928369716e-05
step: 190, loss: 6.951016985112801e-05
step: 200, loss: 0.0007561438833363354
step: 210, loss: 0.00027765092090703547
step: 220, loss: 2.4794975615805015e-05
step: 230, loss: 0.00313349231146276

```

```

/usr/local/lib/python3.7/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=0.9821826280623607, f1=0.9766925638179801,
best_f1=0.9830890642615557
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, 152.54it/s]
load_f1 = 0.9830124575311437
real_f1 = 0.9830124575311437
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, 140.53it/s]

```

4.3.2 DBLP-GoogleScholar - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-GoogleScholar \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm 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.")
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: 0.7186697721481323
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5156172513961792
step: 20, loss: 0.25979161262512207
step: 30, loss: 0.38144221901893616
step: 40, loss: 0.28710585832595825
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.511292576789856
step: 60, loss: 0.1166134625673294
step: 70, loss: 0.19768735766410828
step: 80, loss: 0.17473049461841583
step: 90, loss: 0.1653192788362503
step: 100, loss: 0.2243642657995224
step: 110, loss: 0.10198476165533066
step: 120, loss: 0.18705861270427704
step: 130, loss: 0.08398352563381195
step: 140, loss: 0.24641478061676025
step: 150, loss: 0.14296504855155945
step: 160, loss: 0.13994242250919342
step: 170, loss: 0.09735017269849777
step: 180, loss: 0.03877440094947815
step: 190, loss: 0.11844144016504288
step: 200, loss: 0.028072191402316093

```

```

step: 210, loss: 0.03313669189810753
step: 220, loss: 0.04566288739442825
step: 230, loss: 0.2425198256969452
step: 240, loss: 0.040043026208877563
step: 250, loss: 0.049345508217811584
step: 260, loss: 0.14318758249282837
step: 270, loss: 0.3042718768119812
step: 280, loss: 0.04559057205915451
step: 290, loss: 0.06732980161905289
step: 300, loss: 0.05162119120359421
step: 310, loss: 0.16715869307518005
step: 320, loss: 0.09941054880619049
step: 330, loss: 0.09510256350040436
step: 340, loss: 0.345378041267395
step: 350, loss: 0.132585346698761
step: 360, loss: 0.01508684828877449
step: 370, loss: 0.0445389486849308
step: 380, loss: 0.3624253273010254
step: 390, loss: 0.03864090144634247
step: 400, loss: 0.033186279237270355
step: 410, loss: 0.2456057220697403
step: 420, loss: 0.040223658084869385
step: 430, loss: 0.021550338715314865
step: 440, loss: 0.018996013328433037
step: 450, loss: 0.02438775822520256
step: 460, loss: 0.01661464013159275
step: 470, loss: 0.03694823384284973
step: 480, loss: 0.08691392093896866
step: 490, loss: 0.03938207030296326
step: 500, loss: 0.07337244600057602
step: 510, loss: 0.06628162413835526
step: 520, loss: 0.08217322081327438
step: 530, loss: 0.009968944825232029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9452503509592888, f1=0.9452181987000929,
best_f1=0.9452181987000929
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0728839710354805
step: 10, loss: 0.05367527902126312
step: 20, loss: 0.029296863824129105

```

step: 30, loss: 0.08331894129514694
step: 40, loss: 0.03900160267949104
step: 50, loss: 0.040699660778045654
step: 60, loss: 0.052735839039087296
step: 70, loss: 0.05569510906934738
step: 80, loss: 0.04106895625591278
step: 90, loss: 0.03602186590433121
step: 100, loss: 0.10720524191856384
step: 110, loss: 0.018551629036664963
step: 120, loss: 0.11196035146713257
step: 130, loss: 0.007429295219480991
step: 140, loss: 0.03879259526729584
step: 150, loss: 0.029206791892647743
step: 160, loss: 0.025131039321422577
step: 170, loss: 0.02091952972114086
step: 180, loss: 0.031926173716783524
step: 190, loss: 0.020240671932697296
step: 200, loss: 0.11266077309846878
step: 210, loss: 0.025853481143712997
step: 220, loss: 0.003197565209120512
step: 230, loss: 0.21367846429347992
step: 240, loss: 0.10220682621002197
step: 250, loss: 0.02062767557799816
step: 260, loss: 0.1348968744277954
step: 270, loss: 0.023722916841506958
step: 280, loss: 0.05131791532039642
step: 290, loss: 0.1146778017282486
step: 300, loss: 0.07346880435943604
step: 310, loss: 0.05308642238378525
step: 320, loss: 0.025272339582443237
step: 330, loss: 0.0708639994263649
step: 340, loss: 0.0996766984462738
step: 350, loss: 0.005001731216907501
step: 360, loss: 0.1277955025434494
step: 370, loss: 0.004724243190139532
step: 380, loss: 0.14301468431949615
step: 390, loss: 0.007730184122920036
step: 400, loss: 0.0966104120016098
step: 410, loss: 0.020937180146574974
step: 420, loss: 0.058275289833545685
step: 430, loss: 0.2628032863140106
step: 440, loss: 0.02048531360924244
step: 450, loss: 0.09366129338741302
step: 460, loss: 0.023337604478001595
step: 470, loss: 0.015041862614452839
step: 480, loss: 0.022633017972111702
step: 490, loss: 0.05327963829040527
step: 500, loss: 0.007930624298751354

```

step: 510, loss: 0.03545545041561127
step: 520, loss: 0.37246790528297424
step: 530, loss: 0.06796347349882126
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9499298081422555, 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: 0.14856237173080444
step: 10, loss: 0.056819796562194824
step: 20, loss: 0.009262927807867527
step: 30, loss: 0.1682913899421692
step: 40, loss: 0.06299201399087906
step: 50, loss: 0.13286347687244415
step: 60, loss: 0.06525953114032745
step: 70, loss: 0.002976377494633198
step: 80, loss: 0.12368424981832504
step: 90, loss: 0.17036451399326324
step: 100, loss: 0.01181726437062025
step: 110, loss: 0.04893181100487709
step: 120, loss: 0.057220954447984695
step: 130, loss: 0.1289929747581482
step: 140, loss: 0.022363925352692604
step: 150, loss: 0.040319327265024185
step: 160, loss: 0.013035953044891357
step: 170, loss: 0.014783743768930435
step: 180, loss: 0.016921821981668472
step: 190, loss: 0.0020948939491063356
step: 200, loss: 0.02444242313504219
step: 210, loss: 0.057499371469020844
step: 220, loss: 0.05698798969388008
step: 230, loss: 0.010979389771819115
step: 240, loss: 0.093173086643219
step: 250, loss: 0.0417756587266922
step: 260, loss: 0.07284030318260193
step: 270, loss: 0.0014683741610497236
step: 280, loss: 0.0647696778178215
step: 290, loss: 0.00310420710593462
step: 300, loss: 0.10005242377519608
step: 310, loss: 0.09489548951387405
step: 320, loss: 0.014521252363920212

```

```

step: 330, loss: 0.003633232321590185
step: 340, loss: 0.008929364383220673
step: 350, loss: 0.1528855264186859
step: 360, loss: 0.008396130986511707
step: 370, loss: 0.07368973642587662
step: 380, loss: 0.0071534402668476105
step: 390, loss: 0.07383684813976288
step: 400, loss: 0.08481846749782562
step: 410, loss: 0.0027168062515556812
step: 420, loss: 0.060131218284368515
step: 430, loss: 0.04063705727458
step: 440, loss: 0.3606192469596863
step: 450, loss: 0.040355414152145386
step: 460, loss: 0.16095222532749176
step: 470, loss: 0.009721932001411915
step: 480, loss: 0.030515415593981743
step: 490, loss: 0.0377391017973423
step: 500, loss: 0.016500523313879967
step: 510, loss: 0.03512418642640114
step: 520, loss: 0.06076328083872795
step: 530, loss: 0.004288559779524803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 3: dev_f1=0.9501385041551247, f1=0.951366373320982,
best_f1=0.951366373320982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04929163679480553
step: 10, loss: 0.04455653950572014
step: 20, loss: 0.04072204604744911
step: 30, loss: 0.15664641559123993
step: 40, loss: 0.04437661170959473
step: 50, loss: 0.07034477591514587
step: 60, loss: 0.013014006428420544
step: 70, loss: 0.01731523498892784
step: 80, loss: 0.011205618269741535
step: 90, loss: 0.13779863715171814
step: 100, loss: 0.005781778134405613
step: 110, loss: 0.17936637997627258
step: 120, loss: 0.015725258737802505
step: 130, loss: 0.04059024527668953
step: 140, loss: 0.02781819924712181

```

```

step: 150, loss: 0.008327978663146496
step: 160, loss: 0.041878070682287216
step: 170, loss: 0.04673284664750099
step: 180, loss: 0.07625800371170044
step: 190, loss: 0.04283120110630989
step: 200, loss: 0.09846092015504837
step: 210, loss: 0.004223169758915901
step: 220, loss: 0.008624343201518059
step: 230, loss: 0.032379958778619766
step: 240, loss: 0.06062325835227966
step: 250, loss: 0.18959355354309082
step: 260, loss: 0.0029485097620636225
step: 270, loss: 0.126139298081398
step: 280, loss: 0.03132721036672592
step: 290, loss: 0.03482842817902565
step: 300, loss: 0.001493728719651699
step: 310, loss: 0.0016801775200292468
step: 320, loss: 0.08124915510416031
step: 330, loss: 0.017644796520471573
step: 340, loss: 0.025318874046206474
step: 350, loss: 0.08664069324731827
step: 360, loss: 0.021171629428863525
step: 370, loss: 0.006911568809300661
step: 380, loss: 0.002754536457359791
step: 390, loss: 0.0004016074526589364
step: 400, loss: 0.010257428511977196
step: 410, loss: 0.0296626016497612
step: 420, loss: 0.011555333621799946
step: 430, loss: 0.02748706378042698
step: 440, loss: 0.012145514599978924
step: 450, loss: 0.033954016864299774
step: 460, loss: 0.03644968196749687
step: 470, loss: 0.0046567851677536964
step: 480, loss: 0.07802152633666992
step: 490, loss: 0.10592625290155411
step: 500, loss: 0.03594255819916725
step: 510, loss: 0.052993446588516235
step: 520, loss: 0.012577068991959095
step: 530, loss: 0.15238608419895172
/usr/local/lib/python3.7/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=0.9490592014685636, f1=0.9461573861021628,
best_f1=0.951366373320982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004993739537894726

```

step: 10, loss: 0.049362946301698685
step: 20, loss: 0.10122237354516983
step: 30, loss: 0.04879694804549217
step: 40, loss: 0.0016288121696561575
step: 50, loss: 0.020161932334303856
step: 60, loss: 0.014983843080699444
step: 70, loss: 0.0075163692235946655
step: 80, loss: 0.022807596251368523
step: 90, loss: 0.1065741702914238
step: 100, loss: 0.0936012789607048
step: 110, loss: 0.006750868633389473
step: 120, loss: 0.06772956252098083
step: 130, loss: 0.005611067172139883
step: 140, loss: 0.008888891898095608
step: 150, loss: 0.0017271816032007337
step: 160, loss: 0.0077537307515740395
step: 170, loss: 0.10967505723237991
step: 180, loss: 0.05921241641044617
step: 190, loss: 0.010173737071454525
step: 200, loss: 0.004942539148032665
step: 210, loss: 0.001361656584776938
step: 220, loss: 0.006341190077364445
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 0.00896339863538742
step: 240, loss: 0.017750758677721024
step: 250, loss: 0.12332887947559357
step: 260, loss: 0.0011010190937668085
step: 270, loss: 0.04041225090622902
step: 280, loss: 0.004511129576712847
step: 290, loss: 0.00384151772595942
step: 300, loss: 0.17993468046188354
step: 310, loss: 0.023231089115142822
step: 320, loss: 0.024907035753130913
step: 330, loss: 0.010922976769506931
step: 340, loss: 0.01985847018659115
step: 350, loss: 0.001481160637922585
step: 360, loss: 0.0002863955742213875
step: 370, loss: 0.006255962885916233
step: 380, loss: 0.006840900052338839
step: 390, loss: 0.014679906889796257
step: 400, loss: 0.007993748411536217
step: 410, loss: 0.07332406938076019
step: 420, loss: 0.21450093388557434
step: 430, loss: 0.05789467319846153
step: 440, loss: 0.0050658066757023335
step: 450, loss: 0.009566270746290684
step: 460, loss: 0.0038714748807251453
step: 470, loss: 0.013844083063304424

```

step: 480, loss: 0.022553827613592148
step: 490, loss: 0.01871504634618759
step: 500, loss: 0.08857472985982895
step: 510, loss: 0.020053764805197716
step: 520, loss: 0.05339960381388664
step: 530, loss: 0.04480046033859253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9502093997208004, f1=0.9418931583880038,
best_f1=0.9418931583880038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040115129202604294
step: 10, loss: 0.0009545101202093065
step: 20, loss: 0.030547529458999634
step: 30, loss: 0.0004286256153136492
step: 40, loss: 0.0005472956690937281
step: 50, loss: 0.0006634535966441035
step: 60, loss: 0.1471141278743744
step: 70, loss: 0.0021546112839132547
step: 80, loss: 0.0013812329852953553
step: 90, loss: 0.021714577451348305
step: 100, loss: 0.055197615176439285
step: 110, loss: 0.023468850180506706
step: 120, loss: 0.025741692632436752
step: 130, loss: 0.0017813779413700104
step: 140, loss: 0.0020510510075837374
step: 150, loss: 0.0050397408194839954
step: 160, loss: 0.07696987688541412
step: 170, loss: 0.009952200576663017
step: 180, loss: 0.0008799529750831425
step: 190, loss: 0.26853257417678833
step: 200, loss: 0.022389434278011322
step: 210, loss: 0.04153130203485489
step: 220, loss: 0.015098904259502888
step: 230, loss: 0.17789801955223083
step: 240, loss: 0.01702563837170601
step: 250, loss: 0.017567310482263565
step: 260, loss: 0.004104210529476404
step: 270, loss: 0.04196016490459442
step: 280, loss: 0.006003312300890684
step: 290, loss: 0.0044576385989785194

```



```
step: 300, loss: 0.030752338469028473
step: 310, loss: 0.05894412472844124
step: 320, loss: 0.0006446465849876404
step: 330, loss: 0.00686636520549655
step: 340, loss: 0.0018841152777895331
step: 350, loss: 0.03507326543331146
step: 360, loss: 0.020382734015583992
step: 370, loss: 0.007607604376971722
step: 380, loss: 0.0014533748617395759
step: 390, loss: 0.0022859598975628614
step: 400, loss: 0.005095794331282377
step: 410, loss: 0.000303612498100847
step: 420, loss: 0.0013042327482253313
step: 430, loss: 0.00286984839476645
step: 440, loss: 0.0011648844229057431
step: 450, loss: 0.18188224732875824
step: 460, loss: 0.000789811194408685
step: 470, loss: 0.021184606477618217
step: 480, loss: 0.014782746322453022
step: 490, loss: 0.007064267992973328
step: 500, loss: 0.00481416005641222
step: 510, loss: 0.07296039909124374
step: 520, loss: 0.0009621770586818457
step: 530, loss: 0.0022573331370949745
/usr/local/lib/python3.7/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=0.943327239488117, f1=0.9413919413919414,
best_f1=0.9418931583880038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003720994107425213
step: 10, loss: 0.017756832763552666
step: 20, loss: 0.006209599785506725
step: 30, loss: 0.039727456867694855
step: 40, loss: 0.0179553534835577
step: 50, loss: 0.07981287688016891
step: 60, loss: 0.005020928103476763
step: 70, loss: 0.0005466224392876029
step: 80, loss: 0.00032694736728444695
step: 90, loss: 0.0006470656371675432
step: 100, loss: 0.07884737849235535
step: 110, loss: 0.007008447777479887
step: 120, loss: 0.0015934989787638187
step: 130, loss: 0.001601456431671977
step: 140, loss: 0.0013957256451249123
step: 150, loss: 0.00047056423500180244
```

```

step: 160, loss: 0.00012316271022427827
step: 170, loss: 0.026984872296452522
step: 180, loss: 0.0016709552146494389
step: 190, loss: 0.010632086545228958
step: 200, loss: 0.0002033345226664096
step: 210, loss: 0.1985187828540802
step: 220, loss: 0.11189783364534378
step: 230, loss: 0.018802642822265625
step: 240, loss: 0.019838407635688782
step: 250, loss: 0.015780989080667496
step: 260, loss: 0.015309706330299377
step: 270, loss: 0.05741686373949051
step: 280, loss: 0.0052827573381364346
step: 290, loss: 0.016624413430690765
step: 300, loss: 0.002817069413140416
step: 310, loss: 0.010978582315146923
step: 320, loss: 0.022837884724140167
step: 330, loss: 0.05588260665535927
step: 340, loss: 0.10342521965503693
step: 350, loss: 0.006847957614809275
step: 360, loss: 0.02550382912158966
step: 370, loss: 0.09800507128238678
step: 380, loss: 0.007936048321425915
step: 390, loss: 0.0383458249270916
step: 400, loss: 0.03720949962735176
step: 410, loss: 0.012868903577327728
step: 420, loss: 0.11762090027332306
step: 430, loss: 0.0037887876387685537
step: 440, loss: 0.012692801654338837
step: 450, loss: 0.0019585187546908855
step: 460, loss: 0.012450517155230045
step: 470, loss: 0.13145475089550018
step: 480, loss: 0.006638452876359224
step: 490, loss: 0.004252612125128508
step: 500, loss: 0.004616973455995321
step: 510, loss: 0.000480328977573663
step: 520, loss: 0.0008913209312595427
step: 530, loss: 0.005312428344041109
/usr/local/lib/python3.7/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=0.949343339587242, f1=0.9467232437529468,
best_f1=0.9418931583880038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009555579745210707
step: 10, loss: 0.012414381839334965

```

step: 20, loss: 0.016270095482468605
step: 30, loss: 0.027198394760489464
step: 40, loss: 0.011335449293255806
step: 50, loss: 0.000533124606590718
step: 60, loss: 0.00016922032227739692
step: 70, loss: 0.021123826503753662
step: 80, loss: 0.005878019146621227
step: 90, loss: 0.0003171338757965714
step: 100, loss: 0.00020608576596714556
step: 110, loss: 0.00021520021255128086
step: 120, loss: 0.0014129242626950145
step: 130, loss: 0.0009133280836977065
step: 140, loss: 9.101925388677046e-05
step: 150, loss: 0.007659888826310635
step: 160, loss: 0.0008262947667390108
step: 170, loss: 0.09633934497833252
step: 180, loss: 0.0006823825533501804
step: 190, loss: 0.006279196124523878
step: 200, loss: 0.0071025132201612
step: 210, loss: 0.06881356984376907
step: 220, loss: 0.010378990322351456
step: 230, loss: 0.03178990259766579
step: 240, loss: 0.1520574688911438
step: 250, loss: 0.0006029398064129055
step: 260, loss: 0.1259463131427765
step: 270, loss: 0.03090379759669304
step: 280, loss: 0.00672715762630105
step: 290, loss: 0.001154472935013473
step: 300, loss: 0.00010067486437037587
step: 310, loss: 0.00029322769842110574
step: 320, loss: 0.004856578540056944
step: 330, loss: 0.00010354100959375501
step: 340, loss: 0.04676276072859764
step: 350, loss: 7.929774437798187e-05
step: 360, loss: 0.03916018456220627
step: 370, loss: 0.12557561695575714
step: 380, loss: 0.0009735468192957342
step: 390, loss: 0.01398477517068386
step: 400, loss: 0.006091280374675989
step: 410, loss: 0.00027387242880649865
step: 420, loss: 0.005147624760866165
step: 430, loss: 0.006325528025627136
step: 440, loss: 0.022533319890499115
step: 450, loss: 0.0021848739124834538
step: 460, loss: 0.024075450375676155
step: 470, loss: 0.06817667186260223
step: 480, loss: 0.024594584479928017
step: 490, loss: 0.010595510713756084

```

step: 500, loss: 0.00389675865881145
step: 510, loss: 0.0029722023755311966
step: 520, loss: 0.0008840116788633168
step: 530, loss: 0.0023290084209293127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.950957496496964, f1=0.948155067725362,
best_f1=0.948155067725362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018395804800093174
step: 10, loss: 0.006884968839585781
step: 20, loss: 0.00044560691458173096
step: 30, loss: 0.09584406763315201
step: 40, loss: 0.004716380964964628
step: 50, loss: 0.0011788918636739254
step: 60, loss: 0.0006746758008375764
step: 70, loss: 0.04156243056058884
step: 80, loss: 0.026495778933167458
step: 90, loss: 0.044967737048864365
step: 100, loss: 0.0003410724748391658
step: 110, loss: 0.045943595468997955
step: 120, loss: 0.04047946259379387
step: 130, loss: 0.0013277448015287519
step: 140, loss: 0.02555975317955017
step: 150, loss: 0.0021383774001151323
step: 160, loss: 0.003002743935212493
step: 170, loss: 0.007467441260814667
step: 180, loss: 0.0007020357879810035
step: 190, loss: 0.02928045205771923
step: 200, loss: 0.0012756079668179154
step: 210, loss: 0.006517719477415085
step: 220, loss: 0.01005559228360653
step: 230, loss: 0.0002640754100866616
step: 240, loss: 0.0010225765872746706
step: 250, loss: 0.00445696571841836
step: 260, loss: 0.004657910205423832
step: 270, loss: 0.0007007712847553194
step: 280, loss: 0.004427049774676561
step: 290, loss: 0.0005777982878498733
step: 300, loss: 0.00012365943985059857
step: 310, loss: 0.10954161733388901

```

```

step: 320, loss: 0.006708599627017975
step: 330, loss: 0.00343659077771008
step: 340, loss: 0.007955136708915234
step: 350, loss: 0.06293107569217682
step: 360, loss: 0.001620192313566804
step: 370, loss: 0.000597266131080687
step: 380, loss: 0.0002287691895617172
step: 390, loss: 0.006374327465891838
step: 400, loss: 0.15543010830879211
step: 410, loss: 0.0014770262641832232
step: 420, loss: 0.00019551785953808576
step: 430, loss: 0.03605268895626068
step: 440, loss: 9.430638601770625e-05
step: 450, loss: 0.0016537255141884089
step: 460, loss: 0.00022488407557830215
step: 470, loss: 9.679768845671788e-05
step: 480, loss: 0.0014874276239424944
step: 490, loss: 0.0018895902903750539
step: 500, loss: 0.00017690325330477208
step: 510, loss: 0.028016822412610054
step: 520, loss: 0.0021277505438774824
step: 530, loss: 0.003718130523338914
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9536733738886289, f1=0.9493908153701968,
best_f1=0.9493908153701968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028266076697036624
step: 10, loss: 0.0017342866631224751
step: 20, loss: 0.0015726133715361357
step: 30, loss: 0.008350162766873837
step: 40, loss: 0.0017480294918641448
step: 50, loss: 0.0005841015954501927
step: 60, loss: 0.005240556783974171
step: 70, loss: 0.014556188136339188
step: 80, loss: 0.0008362915832549334
step: 90, loss: 0.0002663703344296664
step: 100, loss: 0.004395362455397844
step: 110, loss: 0.002392729977145791
step: 120, loss: 0.0005776589387096465
step: 130, loss: 0.005048179067671299

```

```

step: 140, loss: 0.0003056130663026124
step: 150, loss: 0.002123587764799595
step: 160, loss: 0.05686042085289955
step: 170, loss: 0.0014460585080087185
step: 180, loss: 0.002199087990447879
step: 190, loss: 0.0005251780967228115
step: 200, loss: 0.00014453120820689946
step: 210, loss: 0.007291749119758606
step: 220, loss: 0.0010613783961161971
step: 230, loss: 0.0004668436595238745
step: 240, loss: 0.0020433671306818724
step: 250, loss: 0.0047757443971931934
step: 260, loss: 0.031927984207868576
step: 270, loss: 0.002067253924906254
step: 280, loss: 0.0014899771194905043
step: 290, loss: 0.00532515486702323
step: 300, loss: 0.013024128042161465
step: 310, loss: 0.05935129523277283
step: 320, loss: 0.009021350182592869
step: 330, loss: 0.02737860567867756
step: 340, loss: 0.0011516871163621545
step: 350, loss: 0.0010756897972896695
step: 360, loss: 0.0001587097649462521
step: 370, loss: 0.00018681729852687567
step: 380, loss: 0.0012256796471774578
step: 390, loss: 0.00024481580476276577
step: 400, loss: 0.0004264444578438997
step: 410, loss: 0.0013325911713764071
step: 420, loss: 9.14537304197438e-05
step: 430, loss: 0.00018156155420001596
step: 440, loss: 2.728970139287412e-05
step: 450, loss: 0.06463466584682465
step: 460, loss: 0.03766009956598282
step: 470, loss: 0.004086120519787073
step: 480, loss: 0.004315941594541073
step: 490, loss: 0.33074283599853516
step: 500, loss: 0.0007027398096397519
step: 510, loss: 0.00019641210383269936
step: 520, loss: 0.004321776330471039
step: 530, loss: 0.024251392111182213
/usr/local/lib/python3.7/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=0.9527340129749768, f1=0.9500462534690102,
best_f1=0.9493908153701968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 4.393274866743013e-05
step: 10, loss: 0.0029401748906821012
step: 20, loss: 0.001910397200845182
step: 30, loss: 0.00032203015871345997
step: 40, loss: 0.0002245938085252419
step: 50, loss: 4.7194644139381126e-05
step: 60, loss: 0.00031132178264670074
step: 70, loss: 0.004188636317849159
step: 80, loss: 0.00012662408698815852
step: 90, loss: 0.016054069623351097
step: 100, loss: 0.003635668195784092
step: 110, loss: 0.0004784745106007904
step: 120, loss: 0.0005787798436358571
step: 130, loss: 8.042231638683006e-05
step: 140, loss: 0.00436064787209034
step: 150, loss: 0.0016325653996318579
step: 160, loss: 0.00036193456617183983
step: 170, loss: 0.0011971734929829836
step: 180, loss: 5.8375895605422556e-05
step: 190, loss: 0.000278716062894091
step: 200, loss: 0.0002716578310355544
step: 210, loss: 0.0008722644415684044
step: 220, loss: 0.0023116031661629677
step: 230, loss: 0.00020928384037688375
step: 240, loss: 0.0022457619197666645
step: 250, loss: 0.008309361524879932
step: 260, loss: 0.0042053028009831905
step: 270, loss: 0.0008626223425380886
step: 280, loss: 0.0008019193774089217
step: 290, loss: 0.0006227773264981806
step: 300, loss: 0.0002107575855916366
step: 310, loss: 0.0005770234856754541
step: 320, loss: 0.0033266337122768164
step: 330, loss: 5.545601015910506e-05
step: 340, loss: 0.0028658646624535322
step: 350, loss: 0.00011314456060063094
step: 360, loss: 0.001445935806259513
step: 370, loss: 0.002266896888613701
step: 380, loss: 0.0005494026700034738
step: 390, loss: 0.0013366015627980232
step: 400, loss: 6.064004628569819e-05
step: 410, loss: 0.0013745097676292062
step: 420, loss: 0.00342698791064322
step: 430, loss: 0.0011884602718055248
step: 440, loss: 0.00013904666411690414
step: 450, loss: 0.0006043943576514721
step: 460, loss: 0.0015201119240373373
step: 470, loss: 0.00019708961190190166

```

step: 480, loss: 0.0021361750550568104
step: 490, loss: 0.009677624329924583
step: 500, loss: 0.0004101296071894467
step: 510, loss: 0.002451175358146429
step: 520, loss: 0.010579276829957962
step: 530, loss: 0.025176942348480225
/usr/local/lib/python3.7/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=0.9517177344475395, f1=0.9486940298507462,
best_f1=0.9493908153701968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009734601713716984
step: 10, loss: 0.004591046832501888
step: 20, loss: 0.002783786505460739
step: 30, loss: 0.050299957394599915
step: 40, loss: 0.0007208720198832452
step: 50, loss: 0.04960445687174797
step: 60, loss: 0.00315713114105165
step: 70, loss: 0.007280423305928707
step: 80, loss: 0.00032496225321665406
step: 90, loss: 0.014724478125572205
step: 100, loss: 0.017770355567336082
step: 110, loss: 0.0014805422397330403
step: 120, loss: 0.0007866713567636907
step: 130, loss: 0.0003923462936654687
step: 140, loss: 0.0001553517213324085
step: 150, loss: 0.006444259081035852
step: 160, loss: 0.00047102541429921985
step: 170, loss: 0.00023275107378140092
step: 180, loss: 0.00014894537162035704
step: 190, loss: 0.00028614638722501695
step: 200, loss: 0.002570555778220296
step: 210, loss: 3.7821042496943846e-05
step: 220, loss: 0.00012149845861131325
step: 230, loss: 0.0075054164044559
step: 240, loss: 8.423566760029644e-05
step: 250, loss: 6.599022162845358e-05
step: 260, loss: 0.00019842527399305254
step: 270, loss: 0.10708639770746231
step: 280, loss: 0.00021171485423110425
step: 290, loss: 0.0030921674333512783
step: 300, loss: 0.0027320515364408493
step: 310, loss: 0.0025366186164319515
step: 320, loss: 0.12816539406776428
step: 330, loss: 0.009902119636535645

```



```

step: 340, loss: 0.002193756867200136
step: 350, loss: 0.0005211834795773029
step: 360, loss: 0.0001976270432351157
step: 370, loss: 0.0011203724425286055
step: 380, loss: 0.0003708887961693108
step: 390, loss: 0.00792939867824316
step: 400, loss: 0.00011885249841725454
step: 410, loss: 0.0009284812258556485
step: 420, loss: 0.016748521476984024
step: 430, loss: 0.003723152680322528
step: 440, loss: 0.00020741995831485838
step: 450, loss: 0.006459229625761509
step: 460, loss: 9.01501189218834e-05
step: 470, loss: 0.004850904457271099
step: 480, loss: 0.0006734266644343734
step: 490, loss: 0.005615759640932083
step: 500, loss: 0.00018195369921159
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 510, loss: 0.015111152082681656
step: 520, loss: 0.003584186313673854
step: 530, loss: 0.0006650544237345457
/usr/local/lib/python3.7/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=0.9534450651769087, f1=0.9497206703910613,
best_f1=0.9493908153701968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004823190101888031
step: 10, loss: 0.0026226763147860765
step: 20, loss: 0.00044997496297582984
step: 30, loss: 6.753360503353179e-05
step: 40, loss: 0.02288222685456276
step: 50, loss: 0.005218889098614454
step: 60, loss: 5.100550697534345e-05
step: 70, loss: 0.005770410876721144
step: 80, loss: 0.009918613359332085
step: 90, loss: 0.0008361660875380039
step: 100, loss: 0.007547697052359581
step: 110, loss: 0.0001656036329222843
step: 120, loss: 0.0006807176396250725
step: 130, loss: 0.0007183275884017348
step: 140, loss: 0.00033047713804990053
step: 150, loss: 0.001970995217561722
step: 160, loss: 0.0008009465527720749
step: 170, loss: 0.007962852716445923
step: 180, loss: 0.0002440608514007181

```

```

step: 190, loss: 0.0014092852361500263
step: 200, loss: 6.560356268892065e-05
step: 210, loss: 3.753467899514362e-05
step: 220, loss: 0.0033950544893741608
step: 230, loss: 0.002885863883420825
step: 240, loss: 0.0013141089584678411
step: 250, loss: 0.0004045399837195873
step: 260, loss: 0.0026730208192020655
step: 270, loss: 0.0026972428895533085
step: 280, loss: 0.00019460092880763113
step: 290, loss: 0.0008355818572454154
step: 300, loss: 0.00015906688349787146
step: 310, loss: 0.0004807894874829799
step: 320, loss: 0.00013172157923690975
step: 330, loss: 0.0008284385548904538
step: 340, loss: 0.0028141250368207693
step: 350, loss: 0.0018668994307518005
step: 360, loss: 0.09380839765071869
step: 370, loss: 0.000970478926319629
step: 380, loss: 0.0014399818610399961
step: 390, loss: 0.00028423217008821666
step: 400, loss: 0.001042571384459734
step: 410, loss: 0.0008972577634267509
step: 420, loss: 0.00022799550788477063
step: 430, loss: 0.00023205050092656165
step: 440, loss: 7.930442370707169e-05
step: 450, loss: 0.004111744463443756
step: 460, loss: 0.01666153408586979
step: 470, loss: 0.0016201059333980083
step: 480, loss: 4.706607069238089e-05
step: 490, loss: 0.0001261645375052467
step: 500, loss: 0.0005920275580137968
step: 510, loss: 0.00016254324873443693
step: 520, loss: 0.00026274428819306195
step: 530, loss: 0.0002956082171294838
/usr/local/lib/python3.7/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=0.9535747446610957, f1=0.9474662947466295,
best_f1=0.9493908153701968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019063655054196715
step: 10, loss: 0.0008269681711681187
step: 20, loss: 0.00024263662635348737
step: 30, loss: 0.0037653185427188873
step: 40, loss: 0.0006924443878233433

```

step: 50, loss: 0.01514937449246645
step: 60, loss: 0.0001864217920228839
step: 70, loss: 0.010250339284539223
step: 80, loss: 8.5297848272603e-05
step: 90, loss: 0.00013689373736269772
step: 100, loss: 0.00013540852523874491
step: 110, loss: 0.0007516042096540332
step: 120, loss: 3.558259049896151e-05
step: 130, loss: 0.00028638209914788604
step: 140, loss: 0.0014165503671392798
step: 150, loss: 0.0013300218852236867
step: 160, loss: 0.0012293662875890732
step: 170, loss: 0.0008185041951946914
step: 180, loss: 0.0005195912672206759
step: 190, loss: 0.00014139499398879707
step: 200, loss: 0.0002495139779057354
step: 210, loss: 0.00048392178723588586
step: 220, loss: 0.00021012221986893564
step: 230, loss: 0.0016748150810599327
step: 240, loss: 0.025914231315255165
step: 250, loss: 0.0016916388412937522
step: 260, loss: 0.0003636789333540946
step: 270, loss: 0.004683308303356171
step: 280, loss: 0.00046365338494069874
step: 290, loss: 0.00011756798630813137
step: 300, loss: 2.824751027219463e-05
step: 310, loss: 0.0004482590011321008
step: 320, loss: 0.003066930454224348
step: 330, loss: 0.013497261330485344
step: 340, loss: 0.000773206411395222
step: 350, loss: 0.00016319150745403022
step: 360, loss: 6.634760211454704e-05
step: 370, loss: 0.01624356396496296
step: 380, loss: 0.000995234353467822
step: 390, loss: 0.0009967545047402382
step: 400, loss: 0.0010780763113871217
step: 410, loss: 0.00015483440074604005
step: 420, loss: 0.0003813181247096509
step: 430, loss: 0.002642435720190406
step: 440, loss: 0.06133534759283066
step: 450, loss: 0.0008431157912127674
step: 460, loss: 0.10499491542577744
step: 470, loss: 4.301925582694821e-05
step: 480, loss: 0.001784725347533822
step: 490, loss: 0.000341972045134753
step: 500, loss: 0.001607418991625309
step: 510, loss: 0.058720678091049194
step: 520, loss: 0.0006037210114300251

```

step: 530, loss: 0.0021058940328657627
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9543568464730291, f1=0.9488243430152145,
best_f1=0.9488243430152145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010063523222925141
step: 10, loss: 0.0035089568700641394
step: 20, loss: 0.0019077819306403399
step: 30, loss: 0.01093259546905756
step: 40, loss: 0.00042408343870192766
step: 50, loss: 0.0020393594168126583
step: 60, loss: 0.0007806257344782352
step: 70, loss: 6.470030348282307e-05
step: 80, loss: 0.00023052070173434913
step: 90, loss: 0.0002498168032616377
step: 100, loss: 0.00010030041448771954
step: 110, loss: 0.0013643479906022549
step: 120, loss: 0.0001629816833883524
step: 130, loss: 0.0010991563322022557
step: 140, loss: 0.00014943112910259515
step: 150, loss: 0.0018703944515436888
step: 160, loss: 0.0003938089939765632
step: 170, loss: 0.00040263403207063675
step: 180, loss: 0.0018733382457867265
step: 190, loss: 0.002126239938661456
step: 200, loss: 0.0022564877290278673
step: 210, loss: 0.0003660871589090675
step: 220, loss: 0.00011323855869704857
step: 230, loss: 0.0008492093184031546
step: 240, loss: 0.00023971179325599223
step: 250, loss: 8.407839777646586e-05
step: 260, loss: 3.008075145771727e-05
step: 270, loss: 0.0003727961448021233
step: 280, loss: 0.00011511080083437264
step: 290, loss: 0.0009109697421081364
step: 300, loss: 0.00033847568556666374
step: 310, loss: 0.012119665741920471
step: 320, loss: 9.62048870860599e-05
step: 330, loss: 6.342729466268793e-05
step: 340, loss: 0.00014133480726741254

```

```

step: 350, loss: 0.0013350986409932375
step: 360, loss: 0.0006063790642656386
step: 370, loss: 0.0005771411815658212
step: 380, loss: 0.0002512536302674562
step: 390, loss: 0.012270198203623295
step: 400, loss: 0.04475509002804756
step: 410, loss: 5.7733108405955136e-05
step: 420, loss: 0.00033869006438180804
step: 430, loss: 6.259030487854034e-05
step: 440, loss: 0.00027851047343574464
step: 450, loss: 0.00193863979075104
step: 460, loss: 0.004614116158336401
step: 470, loss: 0.00025605951668694615
step: 480, loss: 0.019584372639656067
step: 490, loss: 0.00027057831175625324
step: 500, loss: 0.001859538839198649
step: 510, loss: 8.695335418451577e-05
step: 520, loss: 0.0017187736229971051
step: 530, loss: 0.0006718551157973707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9559925093632959, f1=0.9504483246814535,
best_f1=0.9504483246814535
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, 171.46it/s]
load_f1 = 0.956884561891516
real_f1 = 0.9515828677839852
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:29, 147.84it/s]

```