

BDirty_50_3_5_distilbert

August 10, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 22.35 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 103.6 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 61.5 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 50.3 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.6 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 67.7 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 49.8 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 13.16 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
|                                     | 10.4 MB 5.0 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 74.7 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 6.8 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 6.0 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 23.9 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 15.6 MB/s
Requirement already satisfied: scipy>=0.18.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.7.3)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: 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)
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
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 19.6 MB/s
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 23.2 MB/s
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 plac<1.2.0,>=0.9.6
  Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.7/dist-
packages (from spacy==2.2.3->-r requirements.txt (line 4)) (57.4.0)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 19.8 MB/s
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: 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: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Collecting sacremoses

```

```

    Downloading sacremoses-0.0.53.tar.gz (880 kB)
      | 880 kB 22.9 MB/s
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
    | 3.0 MB 24.6 MB/s
Requirement already satisfied: importlib-metadata>=0.20 in /usr/local/lib/python3.7/dist-packages (from catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.1.1)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: urllib3!=1.25.0,!<1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /usr/local/lib/python3.7/dist-packages (from packaging->transformers==3.1.0->-r requirements.txt (line 11)) (3.0.9)
Requirement already satisfied: click in /usr/local/lib/python3.7/dist-packages (from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (7.1.2)
Requirement already satisfied: joblib in /usr/local/lib/python3.7/dist-packages (from sacremoses->transformers==3.1.0->-r requirements.txt (line 11)) (1.1.0)
Requirement already satisfied: threadpoolctl>=2.0.0 in /usr/local/lib/python3.7/dist-packages (from scikit-learn->sklearn==0.0->-r requirements.txt (line 6)) (3.1.0)
Building wheels for collected packages: nltk, sacremoses
  Building wheel for nltk (setup.py) ... done
    Created wheel for nltk: filename=nltk-3.4.5-py3-none-any.whl size=1449924 sha256=4cf0c206feb6d50a77af56e9b5a446ae433bb6ad2c87f0c7c9064dd0faf984fa
    Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b281f323842bbda9bd819eb9a3ead
  Building wheel for sacremoses (setup.py) ... done
    Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl

```

```
size=895260
sha256=97e87d59f0006147915dbfd49813ce7e9b1431ec2c94451122000b226156792f
  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.
```

```
[ ]: True
```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (138/138), done.
remote: Total 10085 (delta 98), reused 131 (delta 59), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 22.75 MiB/s, done.
Resolving deltas: 100% (6902/6902), 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-2knhk0ng
Created temporary directory: /tmp/pip-req-tracker-igy6mrfr
Initialized build tracking at /tmp/pip-req-tracker-igy6mrfr
Created build tracker: /tmp/pip-req-tracker-igy6mrfr
Entered build tracker: /tmp/pip-req-tracker-igy6mrfr
```

```

Created temporary directory: /tmp/pip-install-iyft7uoa
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-099ra37p
  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-
igy6mrfm'
  Running setup.py (path:/tmp/pip-req-build-099ra37p/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-he8ho4h8
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-he8ho4h8/apex.egg-info
  writing /tmp/pip-pip-egg-info-he8ho4h8/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-he8ho4h8/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-he8ho4h8/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-he8ho4h8/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-he8ho4h8/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-099ra37p 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-igy6mrfm'
Created temporary directory: /tmp/pip-unpack-5aq00gai
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-33yv__vg
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-33yv__vg
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''''/tmp/pip-req-build-099ra37p/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-099ra37p/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-33yv_vg
```

```
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/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/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/transformer
copying apex/transformer/__init__.py -> build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/microbatches.py -> build/lib/apex/transformer
copying apex/transformer/log_util.py -> build/lib/apex/transformer
copying apex/transformer/utils.py -> build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
creating build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/opt.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/compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
```

```

copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/frontend.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/_amp_state.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
creating build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_sgd.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_adagrad.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
creating 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
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
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/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/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/peer_memory
copying apex/contrib/peer_memory/__init__.py ->
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_memory.py ->

```

```

build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/__init__.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
  copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/__init__.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/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/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
  copying apex/contrib/multihead_attn/encdec_multihead_attn_func.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/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  creating build/lib/apex/contrib/bottleneck
  copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
  copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck

```

```

    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    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/transducer
    copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    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/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/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/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/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/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels

```

```

copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
creating build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/mappings.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/memory.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/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
creating build/lib/apex/transformer/testing
copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/standalone_bert.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
copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
creating build/lib/apex/transformer/functional
copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
creating build/lib/apex/transformer/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/_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
    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
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    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/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/__init__.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/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
    creating build/lib/apex/amp/lists
    copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
    installing to build/bdist.linux-x86_64/wheel
    running install
    running install_lib
    creating build/bdist.linux-x86_64
    creating build/bdist.linux-x86_64/wheel
    creating build/bdist.linux-x86_64/wheel/apex
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    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
    creating 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/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__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
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.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/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
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
    creating 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_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
    copying build/lib/apex/contrib/optimizers/fused_lamb.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/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    creating 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/self_multihead_attn.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/fast_self_multihead_attn_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
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_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/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating 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/bottleneck_module_test.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/bottleneck.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/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.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/__init__.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/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/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    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/transducer
    copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    creating 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
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    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/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/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/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/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/__init__.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/__init__.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
  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/memory.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/cross_entropy.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/data.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/enums.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/distributed_test_base.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_bert.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/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
  creating build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/functional/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
  copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/amp/__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
    copying build/lib/apex/transformer/utils.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/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying
build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py
-> build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/utils.py ->
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
    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/amp
    copying build/lib/apex/amp/__init__.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/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
  copying build/lib/apex/amp/compat.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/__init__.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/functional_overrides.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/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/scaler.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/frontend.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/_amp_state.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp
  creating 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_sgd.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_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__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
  copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  creating build/bdist.linux-x86_64/wheel/apex/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/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm.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/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/LARC.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/mlp
  copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
  copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
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-33yv__vg/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=298114
sha256=9eb6ca75e646d768100dd2dfc90e354e8788b0a18a79a8d8e5ae709898c8a260
Stored in directory: /tmp/pip-ephem-wheel-cache-2knhk0ng/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-igy6mrfr'
/content/ditto

```

```

[ ]: # some issue with colab
!pip install --upgrade "urllib3==1.25.4" awscli

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 5.2 MB/s
Collecting awscli
  Downloading awscli-1.25.46-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 49.1 MB/s
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 70.6 MB/s
Collecting boto3==1.27.46
  Downloading boto3-1.27.46-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 24.9 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 2.7 MB/s
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==1.27.46->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==1.27.46->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-1.25.46 botocore-1.27.46 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

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

```
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
```

```
Collecting torch==1.7.1
```

```
Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
|                                     | 776.8 MB 17 kB/s
```

```
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
```

```
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-packages
(from torch==1.7.1) (1.19.2)
```

```
Installing collected packages: torch
```

```
Attempting uninstall: torch
Found existing installation: torch 1.4.0
```

```

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

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

torchtex 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

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

```

Cloning into 'IC'...
remote: Enumerating objects: 982, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 982 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (982/982), 251.89 MiB | 29.40 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1273/1273), done.

```

0.4 remove e move

```
[ ]: !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

```

```
[ ]: !rm configs.json
!mv -n /content/ditto/IC/datasEsErros/BDirty_50_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 distilbert \
  --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 distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 442/442 [00:00<00:00, 558kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.16MB/s]
Downloading: 100% 268M/268M [00:05<00:00, 52.4MB/s]
step: 0, loss: 0.873595654964447
epoch 1: dev_f1=0.28, f1=0.2680412371134021, best_f1=0.2680412371134021
step: 0, loss: 0.37537258863449097
epoch 2: dev_f1=0.30434782608695654, f1=0.30952380952380953,
best_f1=0.30952380952380953
step: 0, loss: 0.3563050329685211
epoch 3: dev_f1=0.3548387096774194, f1=0.39215686274509803,
best_f1=0.39215686274509803
step: 0, loss: 0.3649289309978485
```

```

epoch 4: dev_f1=0.34782608695652173, f1=0.42622950819672134,
best_f1=0.39215686274509803
step: 0, loss: 0.3070133328437805
epoch 5: dev_f1=0.36666666666666664, f1=0.423076923076923,
best_f1=0.423076923076923
step: 0, loss: 0.3280651271343231
epoch 6: dev_f1=0.40816326530612246, f1=0.4, best_f1=0.4
step: 0, loss: 0.24604769051074982
epoch 7: dev_f1=0.456140350877193, f1=0.48888888888888893,
best_f1=0.48888888888888893
step: 0, loss: 0.3787340819835663
epoch 8: dev_f1=0.46666666666666667, f1=0.32, best_f1=0.32
step: 0, loss: 0.22358474135398865
epoch 9: dev_f1=0.4615384615384615, f1=0.3448275862068965, best_f1=0.32
step: 0, loss: 0.17945130169391632
epoch 10: dev_f1=0.5000000000000001, f1=0.4137931034482759,
best_f1=0.4137931034482759
step: 0, loss: 0.16728444397449493
epoch 11: dev_f1=0.5, f1=0.375, best_f1=0.4137931034482759
step: 0, loss: 0.20076289772987366
epoch 12: dev_f1=0.5, f1=0.4186046511627907, best_f1=0.4137931034482759
step: 0, loss: 0.1051921546459198
epoch 13: dev_f1=0.5128205128205129, f1=0.4242424242424242,
best_f1=0.4242424242424242
step: 0, loss: 0.09873255342245102
epoch 14: dev_f1=0.5517241379310344, f1=0.35714285714285715,
best_f1=0.35714285714285715
step: 0, loss: 0.13178704679012299
epoch 15: dev_f1=0.5517241379310344, f1=0.35714285714285715,
best_f1=0.35714285714285715
mv: cannot stat '*_dev.pt': No such file or directory
Traceback (most recent call last):
  File "matcher.py", line 315, in <module>
    hp.lm, hp.use_gpu, hp.fp16)
  File "matcher.py", line 274, in load_model
    raise ModelNotFoundError(checkpoint)
ditto_light.exceptions.ModelNotFoundError: Model
checkpoints/Structured/Beer/model.pt was not found

```

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 distilbert \
--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 distilbert \
--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: 0.8152996897697449
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4871104061603546
step: 20, loss: 0.5942121148109436
step: 30, loss: 0.48719266057014465
step: 40, loss: 0.3464900553226471
step: 50, loss: 0.2586989104747772
step: 60, loss: 0.29286885261535645
step: 70, loss: 0.19527724385261536
step: 80, loss: 0.1962575763463974
step: 90, loss: 0.036195263266563416
step: 100, loss: 0.08975756913423538
step: 110, loss: 0.03988468274474144
step: 120, loss: 0.05360366404056549
step: 130, loss: 0.012334105558693409
step: 140, loss: 0.06562740355730057
step: 150, loss: 0.03824646398425102
step: 160, loss: 0.12285508215427399
step: 170, loss: 0.013808047398924828
step: 180, loss: 0.01489266101270914
step: 190, loss: 0.01587124541401863
step: 200, loss: 0.017540343105793
step: 210, loss: 0.028517499566078186
step: 220, loss: 0.004101287107914686
step: 230, loss: 0.10306579619646072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9752808988764046, f1=0.9751131221719457,

```



```

best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031998874619603157
step: 10, loss: 0.0006157690659165382
step: 20, loss: 0.007828985340893269
step: 30, loss: 0.003995357546955347
step: 40, loss: 0.02267937734723091
step: 50, loss: 0.01030945498496294
step: 60, loss: 0.005542417988181114
step: 70, loss: 0.041453953832387924
step: 80, loss: 0.0022040503099560738
step: 90, loss: 0.020438143983483315
step: 100, loss: 0.017838556319475174
step: 110, loss: 0.12313208729028702
step: 120, loss: 0.002040400169789791
step: 130, loss: 0.005648147314786911
step: 140, loss: 0.0665128156542778
step: 150, loss: 0.009693280793726444
step: 160, loss: 0.006805002689361572
step: 170, loss: 0.051329076290130615
step: 180, loss: 0.0016716740792617202
step: 190, loss: 0.24262669682502747
step: 200, loss: 0.13974471390247345
step: 210, loss: 0.05315343663096428
step: 220, loss: 0.0016712944488972425
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 0.17564013600349426
/usr/local/lib/python3.7/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=0.9741863075196409, f1=0.9641255605381166,
best_f1=0.9751131221719457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11058806627988815
step: 10, loss: 0.006050719879567623
step: 20, loss: 0.011331837624311447
step: 30, loss: 0.04986482858657837
step: 40, loss: 0.005482900887727737
step: 50, loss: 0.0030465542804449797
step: 60, loss: 0.026376180350780487
step: 70, loss: 0.0009737666696310043
step: 80, loss: 0.020329851657152176
step: 90, loss: 0.004371316637843847
step: 100, loss: 0.0010474848095327616

```

```

step: 110, loss: 0.0015037300763651729
step: 120, loss: 0.05806625261902809
step: 130, loss: 0.013950871303677559
step: 140, loss: 0.0005153858219273388
step: 150, loss: 0.002796849934384227
step: 160, loss: 0.006974015384912491
step: 170, loss: 0.004147918429225683
step: 180, loss: 0.002914456184953451
step: 190, loss: 0.0030107302591204643
step: 200, loss: 0.0018842347199097276
step: 210, loss: 0.027599165216088295
step: 220, loss: 0.007309053558856249
step: 230, loss: 0.14396166801452637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9798657718120806, f1=0.9741282339707535,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002685489598661661
step: 10, loss: 0.004282550886273384
step: 20, loss: 0.0012405897723510861
step: 30, loss: 0.00033366092247888446
step: 40, loss: 0.012571319006383419
step: 50, loss: 0.002384166233241558
step: 60, loss: 0.035471443086862564
step: 70, loss: 0.056293949484825134
step: 80, loss: 0.055543698370456696
step: 90, loss: 0.0065104723908007145
step: 100, loss: 0.0012208909029141068
step: 110, loss: 0.0240009855479002
step: 120, loss: 0.0020024606492370367
step: 130, loss: 0.0006944878259673715
step: 140, loss: 0.0019924677908420563
step: 150, loss: 0.0010870820842683315
step: 160, loss: 0.010262302123010159
step: 170, loss: 0.0032203819137066603
step: 180, loss: 0.06736833602190018
step: 190, loss: 0.004680991172790527
step: 200, loss: 0.023584486916661263
step: 210, loss: 0.0003401515132281929
step: 220, loss: 0.0011571199866011739

```

```

step: 230, loss: 0.0006690450245514512
/usr/local/lib/python3.7/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=0.9719416386083053, f1=0.9772727272727272,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006642738007940352
step: 10, loss: 0.001925805932842195
step: 20, loss: 0.000608846836257726
step: 30, loss: 0.11638135462999344
step: 40, loss: 0.0005712228594347835
step: 50, loss: 0.0012042996240779757
step: 60, loss: 0.0004271980724297464
step: 70, loss: 0.0004121110250707716
step: 80, loss: 0.001210531685501337
step: 90, loss: 0.008158000186085701
step: 100, loss: 0.023631105199456215
step: 110, loss: 0.0007678584661334753
step: 120, loss: 0.08462204039096832
step: 130, loss: 0.03365866094827652
step: 140, loss: 0.0005345487734302878
step: 150, loss: 0.0004891989519819617
step: 160, loss: 0.0012469387147575617
step: 170, loss: 0.0176601130515337
step: 180, loss: 0.0015305610140785575
step: 190, loss: 0.0004705733445007354
step: 200, loss: 0.012625759467482567
step: 210, loss: 0.00029864584212191403
step: 220, loss: 0.0007748572388663888
step: 230, loss: 0.0015418047551065683
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.9763779527559054,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006426255684345961
step: 10, loss: 0.08340919762849808
step: 20, loss: 0.0011899899691343307
step: 30, loss: 0.0008668222581036389
step: 40, loss: 0.030245207250118256
step: 50, loss: 0.003966289572417736
step: 60, loss: 0.012130173854529858

```

```

step: 70, loss: 0.000652856077067554
step: 80, loss: 0.003914876841008663
step: 90, loss: 0.0001247214386239648
step: 100, loss: 0.00018250780703965575
step: 110, loss: 0.00203106552362442
step: 120, loss: 0.000896961078979075
step: 130, loss: 0.008802274242043495
step: 140, loss: 0.2675928771495819
step: 150, loss: 0.0013153816107660532
step: 160, loss: 0.02827761322259903
step: 170, loss: 0.0021364367567002773
step: 180, loss: 0.002894097473472357
step: 190, loss: 0.0028900648467242718
step: 200, loss: 0.0037298076786100864
step: 210, loss: 0.0026237759739160538
step: 220, loss: 0.0011141576105728745
step: 230, loss: 0.0003393642255105078
/usr/local/lib/python3.7/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=0.9738339021615473, f1=0.9703872437357631,
best_f1=0.9741282339707535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010610549943521619
step: 10, loss: 0.0003682870010379702
step: 20, loss: 0.00043136414024047554
step: 30, loss: 0.0003345150616951287
step: 40, loss: 0.01413656584918499
step: 50, loss: 0.0013340625446289778
step: 60, loss: 0.005974013824015856
step: 70, loss: 0.0010903229704126716
step: 80, loss: 8.947359310695902e-05
step: 90, loss: 0.00022398699366021901
step: 100, loss: 0.0012745494022965431
step: 110, loss: 0.0012350253527984023
step: 120, loss: 0.008644024841487408
step: 130, loss: 0.0017053051851689816
step: 140, loss: 0.000425962294684723
step: 150, loss: 0.0002479424874763936
step: 160, loss: 0.0015521064633503556
step: 170, loss: 0.00015354678907897323
step: 180, loss: 0.0025235360953956842
step: 190, loss: 0.0011732929851859808
step: 200, loss: 0.0028992195148020983
step: 210, loss: 9.442020382266492e-05
step: 220, loss: 0.0006367952446453273

```

```

step: 230, loss: 0.014115597121417522
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9819004524886877, 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: 0.04152324050664902
step: 10, loss: 0.003994164988398552
step: 20, loss: 0.005846067797392607
step: 30, loss: 0.00011115238885395229
step: 40, loss: 0.00011450728197814897
step: 50, loss: 0.0006227571284398437
step: 60, loss: 9.273526666220278e-05
step: 70, loss: 0.0001070000507752411
step: 80, loss: 0.003886888734996319
step: 90, loss: 0.0004750200314447284
step: 100, loss: 0.00018324374104849994
step: 110, loss: 0.00023614399833604693
step: 120, loss: 8.468401210848242e-05
step: 130, loss: 0.005419235676527023
step: 140, loss: 0.00010247856698697433
step: 150, loss: 7.644441211596131e-05
step: 160, loss: 0.00010544853284955025
step: 170, loss: 0.06111632660031319
step: 180, loss: 0.00019885833899024874
step: 190, loss: 0.0034166930709034204
step: 200, loss: 0.011096267960965633
step: 210, loss: 0.0008815810433588922
step: 220, loss: 0.00016440378385595977
step: 230, loss: 0.00025832196115516126
/usr/local/lib/python3.7/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=0.9773755656108598, f1=0.9752808988764046,
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: 0.0001807104708859697
step: 10, loss: 0.00021647765242960304
step: 20, loss: 0.00012849581253249198

```

```

step: 30, loss: 0.00010843913332791999
step: 40, loss: 6.577248859684914e-05
step: 50, loss: 0.00014915478823240846
step: 60, loss: 0.00840416643768549
step: 70, loss: 0.0007415562868118286
step: 80, loss: 0.005955463740974665
step: 90, loss: 0.00022860357421450317
step: 100, loss: 0.0005795734468847513
step: 110, loss: 0.00021415534138213843
step: 120, loss: 0.0027424192521721125
step: 130, loss: 0.00021352349722292274
step: 140, loss: 0.00010789283260237426
step: 150, loss: 0.0009504501358605921
step: 160, loss: 0.0003762533306144178
step: 170, loss: 0.00011523673310875893
step: 180, loss: 0.00022782802989240736
step: 190, loss: 7.556284981546924e-05
step: 200, loss: 0.0003092686238233
step: 210, loss: 0.00041016051545739174
step: 220, loss: 0.08984158933162689
step: 230, loss: 0.00020428860443644226
/usr/local/lib/python3.7/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=0.9744160177975528, f1=0.9766925638179801,
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: 0.0001390753168379888
step: 10, loss: 0.00016688746109139174
step: 20, loss: 0.00021972166723571718
step: 30, loss: 0.002169465646147728
step: 40, loss: 9.338157542515546e-05
step: 50, loss: 0.0236054714769125
step: 60, loss: 0.0002536887186579406
step: 70, loss: 0.0001614007051102817
step: 80, loss: 0.00014783443475607783
step: 90, loss: 0.00012656531180255115
step: 100, loss: 0.00255726370960474
step: 110, loss: 0.00012504680489655584
step: 120, loss: 5.333363878889941e-05
step: 130, loss: 0.00010466777894180268
step: 140, loss: 0.0028621498495340347
step: 150, loss: 0.00026823426014743745
step: 160, loss: 0.00026061851531267166
step: 170, loss: 0.0001476742181694135
step: 180, loss: 0.00013387924991548061

```

```

step: 190, loss: 0.00010200191172771156
step: 200, loss: 7.278987322933972e-05
step: 210, loss: 6.501389725599438e-05
step: 220, loss: 0.0005434390623122454
step: 230, loss: 0.00015356854419223964
/usr/local/lib/python3.7/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=0.976271186440678, f1=0.9832402234636871,
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: 0.00011766893294407055
step: 10, loss: 8.767189865466207e-05
step: 20, loss: 8.675358549226075e-05
step: 30, loss: 9.890257206279784e-05
step: 40, loss: 8.364875975530595e-05
step: 50, loss: 0.0001393362326780334
step: 60, loss: 7.265821477631107e-05
step: 70, loss: 0.00011392825399525464
step: 80, loss: 0.00011083707795478404
step: 90, loss: 0.00013580283848568797
step: 100, loss: 0.006727524567395449
step: 110, loss: 8.896015060599893e-05
step: 120, loss: 0.0002257622982142493
step: 130, loss: 4.3038577132392675e-05
step: 140, loss: 4.304562389734201e-05
step: 150, loss: 6.401622522389516e-05
step: 160, loss: 5.3821004257770255e-05
step: 170, loss: 0.002182516735047102
step: 180, loss: 0.00868368148803711
step: 190, loss: 0.00010987051064148545
step: 200, loss: 7.44222998037003e-05
step: 210, loss: 7.674592052353546e-05
step: 220, loss: 9.278060315409675e-05
step: 230, loss: 0.0002034402423305437
/usr/local/lib/python3.7/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=0.9762174405436014, f1=0.9798206278026906,
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: 0.00017136633687186986
step: 10, loss: 0.0013294105883687735
step: 20, loss: 9.299241355620325e-05

```

```

step: 30, loss: 6.674182077404112e-05
step: 40, loss: 5.916631926083937e-05
step: 50, loss: 4.204252400086261e-05
step: 60, loss: 0.00041596320807002485
step: 70, loss: 5.622518074233085e-05
step: 80, loss: 5.430188321042806e-05
step: 90, loss: 5.1079830882372335e-05
step: 100, loss: 3.897776696248911e-05
step: 110, loss: 6.104989734012634e-05
step: 120, loss: 5.9257243265165016e-05
step: 130, loss: 0.011561527848243713
step: 140, loss: 0.0001918050111271441
step: 150, loss: 4.5071028580423445e-05
step: 160, loss: 0.00012748739391099662
step: 170, loss: 6.490081432275474e-05
step: 180, loss: 5.583660094998777e-05
step: 190, loss: 4.610665200743824e-05
step: 200, loss: 0.006455123890191317
step: 210, loss: 8.154117676895112e-05
step: 220, loss: 7.70206461311318e-05
step: 230, loss: 5.909715764573775e-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=0.979591836734694, f1=0.9796839729119639,
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: 0.0009448293130844831
step: 10, loss: 0.000363703235052526
step: 20, loss: 0.0003116958832833916
step: 30, loss: 0.004981668200343847
step: 40, loss: 7.376621215371415e-05
step: 50, loss: 5.2761577535420656e-05
step: 60, loss: 0.00012220026110298932
step: 70, loss: 6.624733214266598e-05
step: 80, loss: 3.806763197644614e-05
step: 90, loss: 3.75998679373879e-05
step: 100, loss: 6.063104228815064e-05
step: 110, loss: 8.65704205352813e-05
step: 120, loss: 5.5046573834260926e-05
step: 130, loss: 0.0001070832513505593
step: 140, loss: 6.26663604634814e-05
step: 150, loss: 0.02528276853263378
step: 160, loss: 4.1145827708533034e-05
step: 170, loss: 5.1613438699860126e-05
step: 180, loss: 8.308148971991614e-05

```



```

step: 190, loss: 4.27700397267472e-05
step: 200, loss: 5.428036456578411e-05
step: 210, loss: 7.768144860165194e-05
step: 220, loss: 6.9566594902426e-05
step: 230, loss: 3.37339733960107e-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=0.9807037457434733, f1=0.9797752808988766,
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: 0.0001317234564339742
step: 10, loss: 1.9780778529820964e-05
step: 20, loss: 0.0002146268670912832
step: 30, loss: 7.393305713776499e-05
step: 40, loss: 7.638881652383134e-05
step: 50, loss: 5.112184226163663e-05
step: 60, loss: 2.8467380616348237e-05
step: 70, loss: 5.373840394895524e-05
step: 80, loss: 6.213170127011836e-05
step: 90, loss: 0.00035816075978800654
step: 100, loss: 0.00022577776690013707
step: 110, loss: 7.653466309420764e-05
step: 120, loss: 4.4422587961889803e-05
step: 130, loss: 6.642479274887592e-05
step: 140, loss: 0.00015849490591790527
step: 150, loss: 6.286199641181156e-05
step: 160, loss: 3.027363527507987e-05
step: 170, loss: 3.5465996916173026e-05
step: 180, loss: 9.22491672099568e-05
step: 190, loss: 5.618964132736437e-05
step: 200, loss: 0.00012696365593001246
step: 210, loss: 8.359026105608791e-05
step: 220, loss: 0.00019226701988372952
step: 230, loss: 1.9490349586703815e-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=0.9807037457434733, f1=0.9797752808988766,
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: 0.00012393829820211977
step: 10, loss: 0.000330150913214311
step: 20, loss: 4.8409969167551026e-05

```

```

step: 30, loss: 5.684564166585915e-05
step: 40, loss: 7.135557825677097e-05
step: 50, loss: 7.128016295609996e-05
step: 60, loss: 2.980810313601978e-05
step: 70, loss: 4.771012027049437e-05
step: 80, loss: 3.2601983548374847e-05
step: 90, loss: 2.8854354241047986e-05
step: 100, loss: 8.002683898666874e-05
step: 110, loss: 6.008900891174562e-05
step: 120, loss: 0.0016643817070871592
step: 130, loss: 8.152567897923291e-05
step: 140, loss: 6.316577491816133e-05
step: 150, loss: 0.00017720733012538403
step: 160, loss: 4.662424908019602e-05
step: 170, loss: 4.3785763409687206e-05
step: 180, loss: 9.250477887690067e-05
step: 190, loss: 9.725479321787134e-05
step: 200, loss: 7.493326120311394e-05
step: 210, loss: 8.54944228194654e-05
step: 220, loss: 0.00015084556071087718
step: 230, loss: 8.503449498675764e-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=0.9807037457434733, f1=0.9808773903262092,
best_f1=0.9785794813979707
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 [00:08, 304.83it/s]
load_f1 = 0.9807474518686297
real_f1 = 0.9807037457434733
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 [00:11, 389.34it/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 distilbert \
    --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 distilbert \
    --use_gpu \

```

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

```
Downloading: 100% 442/442 [00:00<00:00, 450kB/s]  
Downloading: 100% 232k/232k [00:00<00:00, 1.96MB/s]  
Downloading: 100% 268M/268M [00:10<00:00, 25.3MB/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: 0.7937536239624023
```

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

```
step: 10, loss: 0.4588336646556854
```

```
step: 20, loss: 0.5021977424621582
```

```
step: 30, loss: 0.4687138497829437
```

step: 40, loss: 0.4287235140800476
step: 50, loss: 0.4034635126590729
step: 60, loss: 0.45445626974105835
step: 70, loss: 0.15007707476615906
step: 80, loss: 0.2721162438392639
step: 90, loss: 0.1945633441209793
step: 100, loss: 0.21662713587284088
step: 110, loss: 0.18209028244018555
step: 120, loss: 0.18716712296009064
step: 130, loss: 0.07586455345153809
step: 140, loss: 0.348683625459671
step: 150, loss: 0.02376355230808258
step: 160, loss: 0.12359218299388885
step: 170, loss: 0.3566456139087677
step: 180, loss: 0.12199602276086807
step: 190, loss: 0.08953523635864258
step: 200, loss: 0.20716147124767303
step: 210, loss: 0.13111388683319092
step: 220, loss: 0.29555749893188477
step: 230, loss: 0.15896648168563843
step: 240, loss: 0.22829920053482056
step: 250, loss: 0.18641036748886108
step: 260, loss: 0.030666371807456017
step: 270, loss: 0.03604856878519058
step: 280, loss: 0.15575011074543
step: 290, loss: 0.06947188824415207
step: 300, loss: 0.10392142087221146
step: 310, loss: 0.12763065099716187
step: 320, loss: 0.16174498200416565
step: 330, loss: 0.11969741433858871
step: 340, loss: 0.1562107652425766
step: 350, loss: 0.13584409654140472
step: 360, loss: 0.13657791912555695
step: 370, loss: 0.17728368937969208
step: 380, loss: 0.274941086769104
step: 390, loss: 0.07986897975206375
step: 400, loss: 0.04992365837097168
step: 410, loss: 0.04950492084026337
step: 420, loss: 0.04599180445075035
step: 430, loss: 0.052230242639780045
step: 440, loss: 0.12763969600200653
step: 450, loss: 0.11782258749008179
step: 460, loss: 0.0258244089782238
step: 470, loss: 0.37617507576942444
step: 480, loss: 0.23796267807483673
step: 490, loss: 0.04958781972527504
step: 500, loss: 0.05731045454740524
step: 510, loss: 0.07412583380937576

```

step: 520, loss: 0.132508784532547
step: 530, loss: 0.091537244617939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9044465468306527, f1=0.8966174368747023,
best_f1=0.8966174368747023
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20000019669532776
step: 10, loss: 0.15424084663391113
step: 20, loss: 0.19389580190181732
step: 30, loss: 0.08379629254341125
step: 40, loss: 0.00850061234086752
step: 50, loss: 0.14822009205818176
step: 60, loss: 0.19396665692329407
step: 70, loss: 0.11832528561353683
step: 80, loss: 0.02440699189901352
step: 90, loss: 0.029785823076963425
step: 100, loss: 0.2148641049861908
step: 110, loss: 0.06401720643043518
step: 120, loss: 0.08905881643295288
step: 130, loss: 0.05253353342413902
step: 140, loss: 0.07540754228830338
step: 150, loss: 0.07185815274715424
step: 160, loss: 0.11011172086000443
step: 170, loss: 0.1469394564628601
step: 180, loss: 0.009132372215390205
step: 190, loss: 0.06012352555990219
step: 200, loss: 0.06718239188194275
step: 210, loss: 0.025449899956583977
step: 220, loss: 0.1772586703300476
step: 230, loss: 0.1619124859571457
step: 240, loss: 0.14509783685207367
step: 250, loss: 0.11919265985488892
step: 260, loss: 0.045063819736242294
step: 270, loss: 0.294723242521286
step: 280, loss: 0.3222232460975647
step: 290, loss: 0.1548495590686798
step: 300, loss: 0.09119296818971634
step: 310, loss: 0.07998400181531906
step: 320, loss: 0.12341257184743881
step: 330, loss: 0.035253752022981644

```

```

step: 340, loss: 0.02431635558605194
step: 350, loss: 0.08149731904268265
step: 360, loss: 0.1235751360654831
step: 370, loss: 0.003461875021457672
step: 380, loss: 0.0765703022480011
step: 390, loss: 0.05202215164899826
step: 400, loss: 0.11425808072090149
step: 410, loss: 0.014592028222978115
step: 420, loss: 0.11870404332876205
step: 430, loss: 0.03755253553390503
step: 440, loss: 0.045598775148391724
step: 450, loss: 0.032189060002565384
step: 460, loss: 0.07183875888586044
step: 470, loss: 0.15236322581768036
step: 480, loss: 0.1643100380897522
step: 490, loss: 0.07501839101314545
step: 500, loss: 0.017896583303809166
step: 510, loss: 0.08294707536697388
step: 520, loss: 0.05455987527966499
step: 530, loss: 0.0880008414387703
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9058713886300094, f1=0.9108726084927672,
best_f1=0.9108726084927672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045695334672927856
step: 10, loss: 0.12628719210624695
step: 20, loss: 0.2402684986591339
step: 30, loss: 0.09019003808498383
step: 40, loss: 0.02563776634633541
step: 50, loss: 0.03756865859031677
step: 60, loss: 0.023840203881263733
step: 70, loss: 0.05889515206217766
step: 80, loss: 0.022957244887948036
step: 90, loss: 0.05410481244325638
step: 100, loss: 0.06712851673364639
step: 110, loss: 0.10956761986017227
step: 120, loss: 0.14100611209869385
step: 130, loss: 0.040102001279592514
step: 140, loss: 0.12192758172750473
step: 150, loss: 0.042209308594465256

```

```

step: 160, loss: 0.039136502891778946
step: 170, loss: 0.046013299375772476
step: 180, loss: 0.009609435684978962
step: 190, loss: 0.04691964387893677
step: 200, loss: 0.014027425087988377
step: 210, loss: 0.07566379010677338
step: 220, loss: 0.03261585161089897
step: 230, loss: 0.13834379613399506
step: 240, loss: 0.014531325548887253
step: 250, loss: 0.03681541606783867
step: 260, loss: 0.004998247139155865
step: 270, loss: 0.0022824397310614586
step: 280, loss: 0.03260188177227974
step: 290, loss: 0.04346639662981033
step: 300, loss: 0.16676932573318481
step: 310, loss: 0.04693612456321716
step: 320, loss: 0.17764703929424286
step: 330, loss: 0.010987061075866222
step: 340, loss: 0.02161208912730217
step: 350, loss: 0.09982609748840332
step: 360, loss: 0.033964406698942184
step: 370, loss: 0.006448112428188324
step: 380, loss: 0.01648949831724167
step: 390, loss: 0.09509578347206116
step: 400, loss: 0.11730358749628067
step: 410, loss: 0.016685012727975845
step: 420, loss: 0.011604449711740017
step: 430, loss: 0.021588847041130066
step: 440, loss: 0.06042845547199249
step: 450, loss: 0.20360733568668365
step: 460, loss: 0.1302584409713745
step: 470, loss: 0.03928332403302193
step: 480, loss: 0.03912327438592911
step: 490, loss: 0.01668696478009224
step: 500, loss: 0.14644847810268402
step: 510, loss: 0.01645280234515667
step: 520, loss: 0.003200054867193103
step: 530, loss: 0.10725472867488861
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9136822773186409, f1=0.9183955739972337,
best_f1=0.9183955739972337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```


An input tensor was not cuda.

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

```
step: 0, loss: 0.003407342592254281
step: 10, loss: 0.07797759771347046
step: 20, loss: 0.04180771857500076
step: 30, loss: 0.010115652345120907
step: 40, loss: 0.002644701860845089
step: 50, loss: 0.009110428392887115
step: 60, loss: 0.0009146020747721195
step: 70, loss: 0.000997699680738151
step: 80, loss: 0.013089070096611977
step: 90, loss: 0.03459057956933975
step: 100, loss: 0.020181894302368164
step: 110, loss: 0.005683236755430698
step: 120, loss: 0.001617261441424489
step: 130, loss: 0.012380815111100674
step: 140, loss: 0.04612606391310692
step: 150, loss: 0.0006014772225171328
step: 160, loss: 0.05598670616745949
step: 170, loss: 0.0012216783361509442
step: 180, loss: 0.030624886974692345
step: 190, loss: 0.0057385265827178955
step: 200, loss: 0.006229551509022713
step: 210, loss: 0.03352658823132515
step: 220, loss: 0.06874728202819824
step: 230, loss: 0.33812496066093445
step: 240, loss: 0.11264704912900925
step: 250, loss: 0.002325653098523617
step: 260, loss: 0.09526997059583664
step: 270, loss: 0.048337727785110474
step: 280, loss: 0.020079074427485466
step: 290, loss: 0.1519218385219574
step: 300, loss: 0.004152858164161444
step: 310, loss: 0.023142552003264427
step: 320, loss: 0.025803379714488983
step: 330, loss: 0.04497300833463669
step: 340, loss: 0.05232677608728409
step: 350, loss: 0.0051507046446204185
step: 360, loss: 0.007342238910496235
step: 370, loss: 0.023362400010228157
step: 380, loss: 0.00416272459551692
step: 390, loss: 0.004118176177144051
step: 400, loss: 0.0397980660200119
step: 410, loss: 0.02922016754746437
step: 420, loss: 0.002429736778140068
step: 430, loss: 0.0010957899503409863
step: 440, loss: 0.0830688327550888
step: 450, loss: 0.03755595535039902
```

```

step: 460, loss: 0.11086832731962204
step: 470, loss: 0.0052111269906163216
step: 480, loss: 0.007921675220131874
step: 490, loss: 0.04898430034518242
step: 500, loss: 0.05080825462937355
step: 510, loss: 0.07676739990711212
step: 520, loss: 0.05751302093267441
step: 530, loss: 0.0016425810754299164
/usr/local/lib/python3.7/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=0.9048723897911833, f1=0.9080727951469902,
best_f1=0.9183955739972337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015833476791158319
step: 10, loss: 0.015267131850123405
step: 20, loss: 0.004119811579585075
step: 30, loss: 0.0016881481278687716
step: 40, loss: 0.00032877956982702017
step: 50, loss: 0.04846052825450897
step: 60, loss: 0.00040982195059768856
step: 70, loss: 0.0009381945710629225
step: 80, loss: 0.0033881429117172956
step: 90, loss: 0.0018601237097755075
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.003805711166933179
step: 110, loss: 0.008319990709424019
step: 120, loss: 0.000733743654564023
step: 130, loss: 0.008499070070683956
step: 140, loss: 0.005733973812311888
step: 150, loss: 0.00695638544857502
step: 160, loss: 0.08277354389429092
step: 170, loss: 0.06397068500518799
step: 180, loss: 0.00224487972445786
step: 190, loss: 0.00042403547558933496
step: 200, loss: 0.0007812971016392112
step: 210, loss: 0.03604146093130112
step: 220, loss: 0.00027214837609790266
step: 230, loss: 0.0019185985438525677
step: 240, loss: 0.014165408909320831
step: 250, loss: 0.0021292183082550764
step: 260, loss: 0.008145719766616821
step: 270, loss: 0.008841303177177906
step: 280, loss: 0.021537505090236664
step: 290, loss: 0.001600085524842143
step: 300, loss: 0.046395666897296906

```

```

step: 310, loss: 0.00047130847815424204
step: 320, loss: 0.00907938089221716
step: 330, loss: 0.031109090894460678
step: 340, loss: 0.005134906619787216
step: 350, loss: 0.026658961549401283
step: 360, loss: 0.24632738530635834
step: 370, loss: 0.021561799570918083
step: 380, loss: 0.09950724244117737
step: 390, loss: 0.012811867520213127
step: 400, loss: 0.04687758907675743
step: 410, loss: 0.008590785786509514
step: 420, loss: 0.01359369046986103
step: 430, loss: 0.006040740758180618
step: 440, loss: 0.006006104405969381
step: 450, loss: 0.07770486176013947
step: 460, loss: 0.001372772385366261
step: 470, loss: 0.0013269077753648162
step: 480, loss: 0.009314210154116154
step: 490, loss: 0.1077570840716362
step: 500, loss: 0.006402262486517429
step: 510, loss: 0.0029871943406760693
step: 520, loss: 0.007273811846971512
step: 530, loss: 0.015032454393804073
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9140733859730609, f1=0.9149532710280374,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012385004665702581
step: 10, loss: 0.0013246549060568213
step: 20, loss: 0.0010485434904694557
step: 30, loss: 0.07311610877513885
step: 40, loss: 0.0017180410213768482
step: 50, loss: 0.0008765393286012113
step: 60, loss: 0.00030544374021701515
step: 70, loss: 0.0008378177881240845
step: 80, loss: 0.00033588497899472713
step: 90, loss: 0.0003646713448688388
step: 100, loss: 0.051737554371356964
step: 110, loss: 0.014534443616867065
step: 120, loss: 0.02104608155786991

```

```

step: 130, loss: 0.00016707126633264124
step: 140, loss: 0.035304222255945206
step: 150, loss: 0.020884675905108452
step: 160, loss: 0.002581567270681262
step: 170, loss: 0.00030837117810733616
step: 180, loss: 0.000734105531591922
step: 190, loss: 0.039378516376018524
step: 200, loss: 0.0006197974435053766
step: 210, loss: 0.000502620474435389
step: 220, loss: 0.0005864938721060753
step: 230, loss: 0.000282397901173681
step: 240, loss: 0.01971885934472084
step: 250, loss: 0.0004224508593324572
step: 260, loss: 0.00022560525394510478
step: 270, loss: 0.001324304728768766
step: 280, loss: 0.019513992592692375
step: 290, loss: 0.002339717000722885
step: 300, loss: 0.00429545808583498
step: 310, loss: 0.035733744502067566
step: 320, loss: 0.026369478553533554
step: 330, loss: 0.0014296191511675715
step: 340, loss: 0.001566240913234651
step: 350, loss: 0.04284098744392395
step: 360, loss: 0.00023114090436138213
step: 370, loss: 0.00363246351480484
step: 380, loss: 0.013401112519204617
step: 390, loss: 0.0666845440864563
step: 400, loss: 0.0005209984956309199
step: 410, loss: 0.00041555307689122856
step: 420, loss: 0.0067097945138812065
step: 430, loss: 0.0023522931151092052
step: 440, loss: 0.00890955701470375
step: 450, loss: 0.0015052130911499262
step: 460, loss: 0.0029133628122508526
step: 470, loss: 0.03028273582458496
step: 480, loss: 0.007942231371998787
step: 490, loss: 0.00034037893055938184
step: 500, loss: 0.027628732845187187
step: 510, loss: 0.0001337891153525561
step: 520, loss: 0.22101974487304688
step: 530, loss: 0.08600551635026932
/usr/local/lib/python3.7/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=0.9043560606060606, f1=0.9041487839771101,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009979892522096634
step: 10, loss: 0.015063080005347729
step: 20, loss: 0.0017362534999847412
step: 30, loss: 0.0009060198208317161
step: 40, loss: 0.0005718685570172966
step: 50, loss: 0.009604424238204956
step: 60, loss: 0.0013839849270880222
step: 70, loss: 0.012036914005875587
step: 80, loss: 0.0013446396915242076
step: 90, loss: 0.0004360948223620653
step: 100, loss: 0.000441286334535107
step: 110, loss: 0.00827583484351635
step: 120, loss: 0.0004300941654946655
step: 130, loss: 0.0003506386710796505
step: 140, loss: 0.000741802214179188
step: 150, loss: 0.0003951411054003984
step: 160, loss: 0.0010598321678116918
step: 170, loss: 0.00022513467411044985
step: 180, loss: 0.01954212412238121
step: 190, loss: 0.004948941525071859
step: 200, loss: 0.0005393432220444083
step: 210, loss: 0.03236575797200203
step: 220, loss: 0.004291029181331396
step: 230, loss: 0.010244966484606266
step: 240, loss: 0.0006094390409998596
step: 250, loss: 0.0002589749055914581
step: 260, loss: 0.0001292562228627503
step: 270, loss: 7.736423140158877e-05
step: 280, loss: 0.0756634846329689
step: 290, loss: 0.007394816260784864
step: 300, loss: 0.000283617788227275
step: 310, loss: 0.011524616740643978
step: 320, loss: 0.0026961280964314938
step: 330, loss: 0.0002614824625197798
step: 340, loss: 0.004015721380710602
step: 350, loss: 0.00390800042077899
step: 360, loss: 0.0014302096096798778
step: 370, loss: 0.006167047191411257
step: 380, loss: 0.0013350211083889008
step: 390, loss: 0.0004073514137417078
step: 400, loss: 0.00042105140164494514
step: 410, loss: 0.003518687328323722
step: 420, loss: 0.000806349387858063
step: 430, loss: 6.380792910931632e-05
step: 440, loss: 0.00010949330317089334
step: 450, loss: 0.01897091045975685
step: 460, loss: 0.018876273185014725
```

```

step: 470, loss: 0.030634980648756027
step: 480, loss: 0.0021353529300540686
step: 490, loss: 0.007791465613991022
step: 500, loss: 0.0018195757875218987
step: 510, loss: 0.026143193244934082
step: 520, loss: 0.0060522970743477345
step: 530, loss: 0.004745843820273876
/usr/local/lib/python3.7/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=0.9117370892018779, f1=0.9053420805998126,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013587518595159054
step: 10, loss: 0.001880176831036806
step: 20, loss: 0.01959872618317604
step: 30, loss: 0.001119015971198678
step: 40, loss: 0.0435996949672699
step: 50, loss: 0.0003027870843652636
step: 60, loss: 0.0011270912364125252
step: 70, loss: 0.002137671923264861
step: 80, loss: 0.0012513455003499985
step: 90, loss: 9.459495777264237e-05
step: 100, loss: 0.0642266571521759
step: 110, loss: 0.04248877242207527
step: 120, loss: 0.02669266238808632
step: 130, loss: 0.0030346461571753025
step: 140, loss: 0.0014421074884012341
step: 150, loss: 0.0011780575150623918
step: 160, loss: 0.07895994931459427
step: 170, loss: 0.005699399393051863
step: 180, loss: 0.0005399345536716282
step: 190, loss: 0.06726285815238953
step: 200, loss: 0.00026652589440345764
step: 210, loss: 0.0005177108687348664
step: 220, loss: 0.001451413962058723
step: 230, loss: 0.0003565099323168397
step: 240, loss: 0.01864626817405224
step: 250, loss: 0.0033196802251040936
step: 260, loss: 0.023279236629605293
step: 270, loss: 0.00016458731261081994
step: 280, loss: 0.00016669303295202553
step: 290, loss: 0.041716307401657104
step: 300, loss: 0.00018653189181350172
step: 310, loss: 0.02683795429766178
step: 320, loss: 0.07475633174180984

```

```

step: 330, loss: 0.0247843936085701
step: 340, loss: 0.001984397880733013
step: 350, loss: 0.00715236458927393
step: 360, loss: 0.0010576866334304214
step: 370, loss: 0.022602247074246407
step: 380, loss: 0.03205543011426926
step: 390, loss: 0.0003570659027900547
step: 400, loss: 0.03059128299355507
step: 410, loss: 0.00936111994087696
step: 420, loss: 0.00028178529464639723
step: 430, loss: 0.002637537196278572
step: 440, loss: 0.0013642044505104423
step: 450, loss: 0.00018845607701223344
step: 460, loss: 0.002223121700808406
step: 470, loss: 0.0008412249735556543
step: 480, loss: 0.0007371058454737067
step: 490, loss: 0.00024619564646854997
step: 500, loss: 0.0018340816022828221
step: 510, loss: 0.000935821735765785
step: 520, loss: 0.0003880196891259402
step: 530, loss: 0.00019023886125069112
/usr/local/lib/python3.7/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=0.9130434782608695, f1=0.9122155132373432,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006410912028513849
step: 10, loss: 0.0007121260860003531
step: 20, loss: 0.001145153772085905
step: 30, loss: 6.612852303078398e-05
step: 40, loss: 0.0003926868666894734
step: 50, loss: 0.04301700368523598
step: 60, loss: 8.183817408280447e-05
step: 70, loss: 0.04087013006210327
step: 80, loss: 0.00021602398192044348
step: 90, loss: 4.803836054634303e-05
step: 100, loss: 5.237013101577759e-05
step: 110, loss: 0.0003153455618303269
step: 120, loss: 6.117224984336644e-05
step: 130, loss: 6.526121433125809e-05
step: 140, loss: 7.913581794127822e-05
step: 150, loss: 5.803919702884741e-05
step: 160, loss: 9.59411117946729e-05
step: 170, loss: 5.362584852264263e-05
step: 180, loss: 0.0003543631755746901

```

```

step: 190, loss: 8.10536730568856e-05
step: 200, loss: 0.05658039450645447
step: 210, loss: 6.736500654369593e-05
step: 220, loss: 0.004725123289972544
step: 230, loss: 0.0002680678153410554
step: 240, loss: 0.0001734679244691506
step: 250, loss: 0.00016319160931743681
step: 260, loss: 0.0009333394118584692
step: 270, loss: 0.0001393549027852714
step: 280, loss: 5.75436424696818e-05
step: 290, loss: 3.45467051374726e-05
step: 300, loss: 0.00026920365053229034
step: 310, loss: 4.102338425582275e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 320, loss: 4.454754161997698e-05
step: 330, loss: 0.0008921822300180793
step: 340, loss: 4.196145891910419e-05
step: 350, loss: 3.057986396015622e-05
step: 360, loss: 0.004162505269050598
step: 370, loss: 8.914141653804109e-05
step: 380, loss: 3.5198616387788206e-05
step: 390, loss: 0.00012912016245536506
step: 400, loss: 0.00023196286929305643
step: 410, loss: 0.0017794107552617788
step: 420, loss: 0.0002700819168239832
step: 430, loss: 4.9011003284249455e-05
step: 440, loss: 6.598856998607516e-05
step: 450, loss: 0.0002873994817491621
step: 460, loss: 0.01216064766049385
step: 470, loss: 0.0002713665599003434
step: 480, loss: 4.631876799976453e-05
step: 490, loss: 7.918698247522116e-05
step: 500, loss: 0.024816079065203667
step: 510, loss: 0.0012405001325532794
step: 520, loss: 0.0003652010054793209
step: 530, loss: 0.00018140033353120089
/usr/local/lib/python3.7/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=0.9119850187265918, f1=0.908833254605574,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005579732824116945
step: 10, loss: 0.00013165888958610594
step: 20, loss: 0.000230919846217148
step: 30, loss: 0.012137265875935555

```


step: 40, loss: 0.0008712242124602199
step: 50, loss: 0.00017064480925910175
step: 60, loss: 0.0004152389010414481
step: 70, loss: 0.0011876983335241675
step: 80, loss: 0.0006514975684694946
step: 90, loss: 0.0007658023969270289
step: 100, loss: 8.458639058517292e-05
step: 110, loss: 0.0012913328828290105
step: 120, loss: 7.574389746878296e-05
step: 130, loss: 5.2355888328747824e-05
step: 140, loss: 0.0043483031913638115
step: 150, loss: 0.00020967950695194304
step: 160, loss: 6.0682647017529234e-05
step: 170, loss: 0.00584443798288703
step: 180, loss: 0.0003046572965104133
step: 190, loss: 0.00012841843999922276
step: 200, loss: 0.0003558429889380932
step: 210, loss: 0.00010942461813101545
step: 220, loss: 4.930757495458238e-05
step: 230, loss: 2.7961079467786476e-05
step: 240, loss: 9.979863534681499e-05
step: 250, loss: 0.0008332607103511691
step: 260, loss: 0.011783895082771778
step: 270, loss: 3.7375535612227395e-05
step: 280, loss: 4.2376112105557695e-05
step: 290, loss: 5.014863927499391e-05
step: 300, loss: 5.637199865304865e-05
step: 310, loss: 0.001165176392532885
step: 320, loss: 0.000962872349191457
step: 330, loss: 5.8078345318790525e-05
step: 340, loss: 4.192495543975383e-05
step: 350, loss: 0.00011088236351497471
step: 360, loss: 4.0433264075545594e-05
step: 370, loss: 0.002266284544020891
step: 380, loss: 3.722822657437064e-05
step: 390, loss: 3.9325041143456474e-05
step: 400, loss: 0.00021759953233413398
step: 410, loss: 0.00395215954631567
step: 420, loss: 0.01897580921649933
step: 430, loss: 8.245916978921741e-05
step: 440, loss: 2.7055881218984723e-05
step: 450, loss: 0.0009916748385876417
step: 460, loss: 0.0014105390291661024
step: 470, loss: 2.4753917386988178e-05
step: 480, loss: 0.028623279184103012
step: 490, loss: 0.001975712366402149
step: 500, loss: 0.01791427470743656
step: 510, loss: 0.008221343159675598

```
step: 520, loss: 0.00012784085993189365
step: 530, loss: 0.00014911542530171573
/usr/local/lib/python3.7/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=0.9093419236079153, f1=0.9104477611940298,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.380951577331871e-05
step: 10, loss: 0.020229319110512733
step: 20, loss: 0.05406193435192108
step: 30, loss: 0.0026301469188183546
step: 40, loss: 0.0003105360083281994
step: 50, loss: 0.005285395309329033
step: 60, loss: 0.001047404482960701
step: 70, loss: 3.5324683267390355e-05
step: 80, loss: 0.0018325302517041564
step: 90, loss: 6.186872633406892e-05
step: 100, loss: 0.001319474191404879
step: 110, loss: 0.00013686023885384202
step: 120, loss: 0.00039058359107002616
step: 130, loss: 0.0007192858611233532
step: 140, loss: 0.00011911540786968544
step: 150, loss: 0.00019661961414385587
step: 160, loss: 0.025627149268984795
step: 170, loss: 0.0003519411548040807
step: 180, loss: 0.00010281516733812168
step: 190, loss: 8.960627019405365e-05
step: 200, loss: 0.0004382619808893651
step: 210, loss: 7.123784598661587e-05
step: 220, loss: 7.419969915645197e-05
step: 230, loss: 0.0016844888450577855
step: 240, loss: 2.1907906557316892e-05
step: 250, loss: 0.000679108954500407
step: 260, loss: 5.5781380069674924e-05
step: 270, loss: 0.00011004177940776572
step: 280, loss: 0.005137044470757246
step: 290, loss: 0.0001735314290272072
step: 300, loss: 0.0027144222985953093
step: 310, loss: 0.00017255821148864925
step: 320, loss: 0.04173628240823746
step: 330, loss: 0.0016202019760385156
step: 340, loss: 3.440411092014983e-05
step: 350, loss: 0.0006109643145464361
step: 360, loss: 4.236921085976064e-05
step: 370, loss: 0.00011223583715036511
```

```

step: 380, loss: 7.857301534386352e-05
step: 390, loss: 0.02649437077343464
step: 400, loss: 9.337646042695269e-05
step: 410, loss: 4.0054528653854504e-05
step: 420, loss: 0.0001838113385019824
step: 430, loss: 0.0381830595433712
step: 440, loss: 0.0004965276457369328
step: 450, loss: 3.123862552456558e-05
step: 460, loss: 0.00024544517509639263
step: 470, loss: 0.00010152222239412367
step: 480, loss: 4.810768587049097e-05
step: 490, loss: 0.0008980510174296796
step: 500, loss: 3.638909765868448e-05
step: 510, loss: 8.443515980616212e-05
step: 520, loss: 2.3575736122438684e-05
step: 530, loss: 1.7311145711573772e-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=0.9137931034482759, f1=0.9200726612170754,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.9350808947347105e-05
step: 10, loss: 3.642324008978903e-05
step: 20, loss: 2.3434202375938185e-05
step: 30, loss: 7.562092650914565e-05
step: 40, loss: 3.7808924389537424e-05
step: 50, loss: 1.672990038059652e-05
step: 60, loss: 5.181227606954053e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.04334494099020958
step: 80, loss: 0.00026246486231684685
step: 90, loss: 0.00020727215451188385
step: 100, loss: 0.009471043944358826
step: 110, loss: 1.5731691746623255e-05
step: 120, loss: 0.001721005653962493
step: 130, loss: 0.0005377482157200575
step: 140, loss: 3.436012048041448e-05
step: 150, loss: 0.00012417681864462793
step: 160, loss: 0.0001467238471377641
step: 170, loss: 6.118890451034531e-05
step: 180, loss: 6.142750498838723e-05
step: 190, loss: 0.0001042009171214886
step: 200, loss: 0.00023185336613096297
step: 210, loss: 0.00017347473476547748
step: 220, loss: 3.481185558484867e-05

```

```

step: 230, loss: 0.000291056843707338
step: 240, loss: 0.0001411190169164911
step: 250, loss: 2.793475505313836e-05
step: 260, loss: 3.800850026891567e-05
step: 270, loss: 3.052211832255125e-05
step: 280, loss: 2.59219050349202e-05
step: 290, loss: 7.172650657594204e-05
step: 300, loss: 0.0027635591104626656
step: 310, loss: 0.0001626437733648345
step: 320, loss: 4.633748540072702e-05
step: 330, loss: 1.5977573639247566e-05
step: 340, loss: 8.338476618519053e-05
step: 350, loss: 0.00015024545427877456
step: 360, loss: 7.355917477980256e-05
step: 370, loss: 2.4317476345459e-05
step: 380, loss: 0.001322917640209198
step: 390, loss: 0.002584730973467231
step: 400, loss: 0.00013189006131142378
step: 410, loss: 5.166531991562806e-05
step: 420, loss: 0.001284308615140617
step: 430, loss: 0.00010890061821555719
step: 440, loss: 0.00048712617717683315
step: 450, loss: 0.0001569199375808239
step: 460, loss: 0.0008781295618973672
step: 470, loss: 8.436635107500479e-05
step: 480, loss: 0.0005720055196434259
step: 490, loss: 2.8906544685014524e-05
step: 500, loss: 0.0005274809082038701
step: 510, loss: 8.596769475843757e-05
step: 520, loss: 0.00041062975651584566
step: 530, loss: 0.0016795233823359013
/usr/local/lib/python3.7/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=0.9136258660508083, f1=0.9182915506035284,
best_f1=0.9149532710280374
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015425877063535154
step: 10, loss: 0.00011672563414322212
step: 20, loss: 0.0005070515326224267
step: 30, loss: 0.0021817851811647415
step: 40, loss: 2.194487751694396e-05
step: 50, loss: 0.000117100054922048
step: 60, loss: 3.603472941904329e-05
step: 70, loss: 5.541511927731335e-05
step: 80, loss: 0.00010198036034125835

```

```
step: 90, loss: 0.00036740288487635553
step: 100, loss: 0.00010826460493262857
step: 110, loss: 3.2789634133223444e-05
step: 120, loss: 4.0365401218878105e-05
step: 130, loss: 0.001216474687680602
step: 140, loss: 0.0023721931502223015
step: 150, loss: 0.00012831062485929579
step: 160, loss: 3.0049139240873046e-05
step: 170, loss: 6.073248732718639e-05
step: 180, loss: 2.6187288312939927e-05
step: 190, loss: 6.249296711757779e-05
step: 200, loss: 2.5520344934193417e-05
step: 210, loss: 1.5854249795665964e-05
step: 220, loss: 1.843231984821614e-05
step: 230, loss: 3.085753269260749e-05
step: 240, loss: 3.4109732951037586e-05
step: 250, loss: 0.0004045348905492574
step: 260, loss: 3.4764754673233256e-05
step: 270, loss: 2.2048883693059906e-05
step: 280, loss: 2.449007297400385e-05
step: 290, loss: 0.00014676153659820557
step: 300, loss: 2.2589108993997797e-05
step: 310, loss: 8.578804408898577e-05
step: 320, loss: 6.58652643323876e-05
step: 330, loss: 4.617220474756323e-05
step: 340, loss: 0.01928340271115303
step: 350, loss: 0.04586052894592285
step: 360, loss: 3.859507341985591e-05
step: 370, loss: 3.1624131224816665e-05
step: 380, loss: 0.00641494570299983
step: 390, loss: 1.316504040005384e-05
step: 400, loss: 1.3157568901078776e-05
step: 410, loss: 1.8242215446662158e-05
step: 420, loss: 4.4012078433297575e-05
step: 430, loss: 1.4822741832176689e-05
step: 440, loss: 1.9192142644897103e-05
step: 450, loss: 1.5724259355920367e-05
step: 460, loss: 1.2688194146903697e-05
step: 470, loss: 0.00028931026463396847
step: 480, loss: 1.9836406863760203e-05
step: 490, loss: 2.086839231196791e-05
step: 500, loss: 2.1237461623968557e-05
step: 510, loss: 4.679283665609546e-05
step: 520, loss: 1.308306946157245e-05
step: 530, loss: 1.885314668470528e-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=0.9133640552995391, f1=0.9165501165501165,  
best_f1=0.9149532710280374  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 1.746379530231934e-05  
step: 10, loss: 0.0011178478598594666  
step: 20, loss: 1.4312377970782109e-05  
step: 30, loss: 2.2861424440634437e-05  
step: 40, loss: 0.00036337930941954255  
step: 50, loss: 0.00010143845429411158  
step: 60, loss: 4.211736813886091e-05  
step: 70, loss: 1.2591341146617197e-05  
step: 80, loss: 2.3997075913939625e-05  
step: 90, loss: 1.11310655483976e-05  
step: 100, loss: 4.8633242840878665e-05  
step: 110, loss: 1.668887671257835e-05  
step: 120, loss: 1.7478485460742377e-05  
step: 130, loss: 6.451335502788424e-05  
step: 140, loss: 0.001508912188000977  
step: 150, loss: 1.564975900691934e-05  
step: 160, loss: 2.8439590096240863e-05  
step: 170, loss: 0.00025564240058884025  
step: 180, loss: 1.57987105922075e-05  
step: 190, loss: 3.695498526212759e-05  
step: 200, loss: 2.1561481844400987e-05  
step: 210, loss: 0.00033056738902814686  
step: 220, loss: 0.02657558023929596  
step: 230, loss: 1.9829214579658583e-05  
step: 240, loss: 1.330287523160223e-05  
step: 250, loss: 2.088354085572064e-05  
step: 260, loss: 9.555299584462773e-06  
step: 270, loss: 7.015021401457489e-05  
step: 280, loss: 3.654893589555286e-05  
step: 290, loss: 1.5649715351173654e-05  
step: 300, loss: 1.378340857627336e-05  
step: 310, loss: 2.5060704501811415e-05  
step: 320, loss: 0.0001768284710124135  
step: 330, loss: 2.6641520889825188e-05  
step: 340, loss: 2.639639023982454e-05  
step: 350, loss: 1.5925310435704887e-05  
step: 360, loss: 2.9821996577084064e-05  
step: 370, loss: 0.00020603301527444273  
step: 380, loss: 1.1157121662108693e-05  
step: 390, loss: 3.69813060387969e-05  
step: 400, loss: 2.8478767490014434e-05  
step: 410, loss: 0.00031175921321846545  
step: 420, loss: 1.3384811609284952e-05
```

```

step: 430, loss: 1.5120693205972202e-05
step: 440, loss: 1.1872374670929275e-05
step: 450, loss: 1.8163998902309686e-05
step: 460, loss: 2.787769517453853e-05
step: 470, loss: 1.2699356375378557e-05
step: 480, loss: 1.3835528079653159e-05
step: 490, loss: 0.00018414293299429119
step: 500, loss: 0.0030385327991098166
step: 510, loss: 0.00015023292507976294
step: 520, loss: 1.922578849189449e-05
step: 530, loss: 0.0003903727629221976
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9158708503865394, f1=0.9172727272727272,
best_f1=0.9172727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.8599654140416533e-05
step: 10, loss: 6.822751311119646e-05
step: 20, loss: 0.0001250630448339507
step: 30, loss: 4.7570742026437074e-05
step: 40, loss: 9.96870658127591e-05
step: 50, loss: 6.182816287036985e-05
step: 60, loss: 9.483182657277212e-05
step: 70, loss: 2.0663272152887657e-05
step: 80, loss: 1.8585093130241148e-05
step: 90, loss: 1.3470416888594627e-05
step: 100, loss: 0.028775904327630997
step: 110, loss: 0.0015747882425785065
step: 120, loss: 5.8771958720171824e-05
step: 130, loss: 3.0264158340287395e-05
step: 140, loss: 1.6182257240870968e-05
step: 150, loss: 0.00031696504447609186
step: 160, loss: 8.462231198791414e-05
step: 170, loss: 2.7210968255531043e-05
step: 180, loss: 0.003537448588758707
step: 190, loss: 2.866048271243926e-05
step: 200, loss: 1.476674424338853e-05
step: 210, loss: 2.4777840735623613e-05
step: 220, loss: 2.264614886371419e-05
step: 230, loss: 0.0021365240681916475
step: 240, loss: 1.695351420494262e-05

```

```

step: 250, loss: 2.7572257749852724e-05
step: 260, loss: 0.0005072229541838169
step: 270, loss: 0.06462006270885468
step: 280, loss: 2.0667417629738338e-05
step: 290, loss: 5.93559343542438e-05
step: 300, loss: 3.675224070320837e-05
step: 310, loss: 3.9121830923249945e-05
step: 320, loss: 1.6551217413507402e-05
step: 330, loss: 4.945716864312999e-05
step: 340, loss: 0.0026647511404007673
step: 350, loss: 0.002524276729673147
step: 360, loss: 0.000164839526405558
step: 370, loss: 1.9776320186792873e-05
step: 380, loss: 1.3727494660997763e-05
step: 390, loss: 0.00015732126485090703
step: 400, loss: 6.191136344568804e-05
step: 410, loss: 9.494894038653001e-05
step: 420, loss: 4.904405795969069e-05
step: 430, loss: 1.4338446817419026e-05
step: 440, loss: 0.0019109423737972975
step: 450, loss: 0.021495964378118515
step: 460, loss: 1.7254866179428063e-05
step: 470, loss: 2.4413273422396742e-05
step: 480, loss: 5.929892722633667e-05
step: 490, loss: 2.978088377858512e-05
step: 500, loss: 2.4894336092984304e-05
step: 510, loss: 2.1873813238926232e-05
step: 520, loss: 0.00024560370366089046
step: 530, loss: 1.0631898476276547e-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=0.9140229885057471, f1=0.9180176007410837,
best_f1=0.9172727272727272
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:23, 241.18it/s]
load_f1 = 0.9129239230064162
real_f1 = 0.9102564102564102
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 [00:10, 432.44it/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 distilbert \
    --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 distilbert \
  --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: 0.8477713465690613
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.07316890358924866
step: 20, loss: 0.3761886656284332
step: 30, loss: 0.3685636520385742
step: 40, loss: 0.514959454536438
step: 50, loss: 0.3097822666168213
step: 60, loss: 0.37267550826072693
step: 70, loss: 0.2265503853559494
step: 80, loss: 0.26343727111816406
step: 90, loss: 0.42117080092430115
step: 100, loss: 0.17088265717029572
step: 110, loss: 0.34565088152885437
step: 120, loss: 0.2644636929035187
step: 130, loss: 0.23972845077514648
step: 140, loss: 0.2537247836589813
step: 150, loss: 0.2746618390083313
step: 160, loss: 0.23549039661884308
step: 170, loss: 0.16917675733566284
step: 180, loss: 0.19576500356197357
step: 190, loss: 0.2296963930130005
step: 200, loss: 0.20483297109603882
step: 210, loss: 0.4558337926864624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4904051172707889, f1=0.50989010989011,
best_f1=0.50989010989011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07757706195116043
step: 10, loss: 0.10550220310688019
step: 20, loss: 0.25119340419769287
step: 30, loss: 0.19839444756507874
step: 40, loss: 0.10838635265827179
step: 50, loss: 0.19147348403930664
step: 60, loss: 0.08866336196660995
step: 70, loss: 0.2539500892162323
step: 80, loss: 0.22207483649253845
step: 90, loss: 0.10338295251131058
step: 100, loss: 0.1607154756784439
step: 110, loss: 0.0829387903213501
step: 120, loss: 0.2006806880235672

```

```

step: 130, loss: 0.25748327374458313
step: 140, loss: 0.2558552026748657
step: 150, loss: 0.31621435284614563
step: 160, loss: 0.1313362568616867
step: 170, loss: 0.2300824522972107
step: 180, loss: 0.1455451250076294
step: 190, loss: 0.17758512496948242
step: 200, loss: 0.17875637114048004
step: 210, loss: 0.3562759757041931
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5725190839694657, f1=0.5371428571428571,
best_f1=0.5371428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1576550453901291
step: 10, loss: 0.3056201636791229
step: 20, loss: 0.3600960373878479
step: 30, loss: 0.058125514537096024
step: 40, loss: 0.17718254029750824
step: 50, loss: 0.22772841155529022
step: 60, loss: 0.1962793469429016
step: 70, loss: 0.09916073828935623
step: 80, loss: 0.17064359784126282
step: 90, loss: 0.16426512598991394
step: 100, loss: 0.0770522877573967
step: 110, loss: 0.23668743669986725
step: 120, loss: 0.09701760113239288
step: 130, loss: 0.1749303936958313
step: 140, loss: 0.20719733834266663
step: 150, loss: 0.26011744141578674
step: 160, loss: 0.11150147020816803
step: 170, loss: 0.20997118949890137
step: 180, loss: 0.20730134844779968
step: 190, loss: 0.24459411203861237
step: 200, loss: 0.16891765594482422
step: 210, loss: 0.21303904056549072
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5759637188208616, f1=0.5525114155251142,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27954819798469543
step: 10, loss: 0.10368301719427109
step: 20, loss: 0.16084207594394684
step: 30, loss: 0.05050864815711975
step: 40, loss: 0.07804493606090546
step: 50, loss: 0.08445927500724792
step: 60, loss: 0.03653683513402939
step: 70, loss: 0.2523983120918274
step: 80, loss: 0.18932946026325226
step: 90, loss: 0.11211222410202026
step: 100, loss: 0.14800430834293365
step: 110, loss: 0.10169261693954468
step: 120, loss: 0.12956677377223969
step: 130, loss: 0.24476240575313568
step: 140, loss: 0.1893230825662613
step: 150, loss: 0.25216612219810486
step: 160, loss: 0.09791813045740128
step: 170, loss: 0.06757678836584091
step: 180, loss: 0.20738807320594788
step: 190, loss: 0.10053670406341553
step: 200, loss: 0.11594165861606598
step: 210, loss: 0.06793022900819778
/usr/local/lib/python3.7/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=0.5754527162977867, f1=0.5599999999999999,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10725284367799759
step: 10, loss: 0.0661528930068016
step: 20, loss: 0.16304312646389008
step: 30, loss: 0.17337827384471893
step: 40, loss: 0.42661672830581665
step: 50, loss: 0.2157445102930069
step: 60, loss: 0.2408440262079239
step: 70, loss: 0.09201895445585251
step: 80, loss: 0.0373750738799572
step: 90, loss: 0.05423008278012276
step: 100, loss: 0.11384832859039307

```

```

step: 110, loss: 0.021596074104309082
step: 120, loss: 0.027956021949648857
step: 130, loss: 0.0925905704498291
step: 140, loss: 0.1478683203458786
step: 150, loss: 0.2105945348739624
step: 160, loss: 0.03714043274521828
step: 170, loss: 0.07371442764997482
step: 180, loss: 0.15109656751155853
step: 190, loss: 0.11338669061660767
step: 200, loss: 0.12929242849349976
step: 210, loss: 0.041556864976882935
/usr/local/lib/python3.7/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=0.5507900677200904, f1=0.5023474178403756,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05941079929471016
step: 10, loss: 0.039484646171331406
step: 20, loss: 0.11501868814229965
step: 30, loss: 0.1647428274154663
step: 40, loss: 0.04630609229207039
step: 50, loss: 0.03078315034508705
step: 60, loss: 0.11568471789360046
step: 70, loss: 0.02063380926847458
step: 80, loss: 0.1946132630109787
step: 90, loss: 0.07556816935539246
step: 100, loss: 0.03544529154896736
step: 110, loss: 0.05020417273044586
step: 120, loss: 0.02146102674305439
step: 130, loss: 0.03196515515446663
step: 140, loss: 0.1418733149766922
step: 150, loss: 0.059287555515766144
step: 160, loss: 0.38563770055770874
step: 170, loss: 0.07098692655563354
step: 180, loss: 0.013040587306022644
step: 190, loss: 0.1172911524772644
step: 200, loss: 0.02939092181622982
step: 210, loss: 0.02887207828462124
/usr/local/lib/python3.7/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=0.5608247422680411, f1=0.5258620689655173,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020212559029459953
step: 10, loss: 0.06838490813970566
step: 20, loss: 0.014275558292865753
step: 30, loss: 0.0041794199496507645
step: 40, loss: 0.052600517868995667
step: 50, loss: 0.07730674743652344
step: 60, loss: 0.25737887620925903
step: 70, loss: 0.03924035280942917
step: 80, loss: 0.018311534076929092
step: 90, loss: 0.009483952075242996
step: 100, loss: 0.020532943308353424
step: 110, loss: 0.2256961315870285
step: 120, loss: 0.14000192284584045
step: 130, loss: 0.008850045502185822
step: 140, loss: 0.004782022442668676
step: 150, loss: 0.02600850537419319
step: 160, loss: 0.0662190243601799
step: 170, loss: 0.052445583045482635
step: 180, loss: 0.007017288822680712
step: 190, loss: 0.05238654837012291
step: 200, loss: 0.01601356826722622
step: 210, loss: 0.04385817050933838
/usr/local/lib/python3.7/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=0.5475728155339807, f1=0.5521235521235521,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028342269361019135
step: 10, loss: 0.013435586355626583
step: 20, loss: 0.10923309624195099
step: 30, loss: 0.004007443320006132
step: 40, loss: 0.014178439974784851
step: 50, loss: 0.022133488208055496
step: 60, loss: 0.15678872168064117
step: 70, loss: 0.012915448285639286
step: 80, loss: 0.02063589170575142
step: 90, loss: 0.04803686961531639
step: 100, loss: 0.014094662852585316
step: 110, loss: 0.001584534882567823
step: 120, loss: 0.01671929843723774
step: 130, loss: 0.015316735953092575
step: 140, loss: 0.0034383947495371103
step: 150, loss: 0.02480699121952057
step: 160, loss: 0.009114308282732964

```

```

step: 170, loss: 0.00723244808614254
step: 180, loss: 0.11721954494714737
step: 190, loss: 0.07955624908208847
step: 200, loss: 0.07925107330083847
step: 210, loss: 0.11436181515455246
/usr/local/lib/python3.7/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=0.5241379310344828, f1=0.5091743119266054,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02299930527806282
step: 10, loss: 0.1158539280295372
step: 20, loss: 0.11164925992488861
step: 30, loss: 0.12322758883237839
step: 40, loss: 0.129150390625
step: 50, loss: 0.004174038767814636
step: 60, loss: 0.043482471257448196
step: 70, loss: 0.03541122376918793
step: 80, loss: 0.022587157785892487
step: 90, loss: 0.1329687237739563
step: 100, loss: 0.03562832996249199
step: 110, loss: 0.0021815297659486532
step: 120, loss: 0.007295048329979181
step: 130, loss: 0.14940626919269562
step: 140, loss: 0.056159187108278275
step: 150, loss: 0.06767489016056061
step: 160, loss: 0.006011580117046833
step: 170, loss: 0.17989055812358856
step: 180, loss: 0.015685098245739937
step: 190, loss: 0.1617317646741867
step: 200, loss: 0.026952389627695084
step: 210, loss: 0.03358963504433632
/usr/local/lib/python3.7/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=0.5351473922902494, f1=0.5137614678899082,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11191105842590332
step: 10, loss: 0.026286981999874115
step: 20, loss: 0.05374931916594505
step: 30, loss: 0.01900455169379711
step: 40, loss: 0.06944003701210022

```



```

step: 50, loss: 0.02558048628270626
step: 60, loss: 0.0020703342743217945
step: 70, loss: 0.001085186144337058
step: 80, loss: 0.0005293729482218623
step: 90, loss: 0.028326204046607018
step: 100, loss: 0.0034421749878674746
step: 110, loss: 0.011871563270688057
step: 120, loss: 0.027364209294319153
step: 130, loss: 0.003121788613498211
step: 140, loss: 0.007156732492148876
step: 150, loss: 0.08932366967201233
step: 160, loss: 0.15797002613544464
step: 170, loss: 0.03301025182008743
step: 180, loss: 0.010494833812117577
step: 190, loss: 0.21676428616046906
step: 200, loss: 0.06318876892328262
step: 210, loss: 0.007013530936092138
/usr/local/lib/python3.7/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=0.5178947368421053, f1=0.5387931034482758,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043333750218153
step: 10, loss: 0.005732567049562931
step: 20, loss: 0.010341677814722061
step: 30, loss: 0.013011268340051174
step: 40, loss: 0.03239366039633751
step: 50, loss: 0.02043999545276165
step: 60, loss: 0.007389376871287823
step: 70, loss: 0.0032263793982565403
step: 80, loss: 0.15753592550754547
step: 90, loss: 0.036413054913282394
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.007095435634255409
step: 110, loss: 0.004019269719719887
step: 120, loss: 0.0023962410632520914
step: 130, loss: 0.05695006623864174
step: 140, loss: 0.010274814441800117
step: 150, loss: 0.027217814698815346
step: 160, loss: 0.00346600404009223
step: 170, loss: 0.15620943903923035
step: 180, loss: 0.045634929090738297
step: 190, loss: 0.026407703757286072
step: 200, loss: 0.05691056698560715
step: 210, loss: 0.007208780385553837

```

```

/usr/local/lib/python3.7/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=0.519774011299435, f1=0.5527831094049904,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008239329792559147
step: 10, loss: 0.005118754226714373
step: 20, loss: 0.019441237673163414
step: 30, loss: 0.10391046851873398
step: 40, loss: 0.004152797162532806
step: 50, loss: 0.047292932868003845
step: 60, loss: 0.012312745675444603
step: 70, loss: 0.012418640777468681
step: 80, loss: 0.01215648464858532
step: 90, loss: 0.01045677438378334
step: 100, loss: 0.0782194584608078
step: 110, loss: 0.0034643877297639847
step: 120, loss: 0.009794597513973713
step: 130, loss: 0.03908814117312431
step: 140, loss: 0.003244026331230998
step: 150, loss: 0.005040797404944897
step: 160, loss: 0.026635071262717247
step: 170, loss: 0.011665581725537777
step: 180, loss: 0.002341249492019415
step: 190, loss: 0.04261763393878937
step: 200, loss: 0.07115932554006577
step: 210, loss: 0.0021374563220888376
/usr/local/lib/python3.7/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=0.5142857142857143, f1=0.5481171548117155,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018712218152359128
step: 10, loss: 0.008273565210402012
step: 20, loss: 0.024846818298101425
step: 30, loss: 0.02947554551064968
step: 40, loss: 0.005314561538398266
step: 50, loss: 0.0114594092592597
step: 60, loss: 0.004319216124713421
step: 70, loss: 0.0714261382818222
step: 80, loss: 0.02036943845450878
step: 90, loss: 0.01207401417195797

```

```

step: 100, loss: 0.17408180236816406
step: 110, loss: 0.00217165844514966
step: 120, loss: 0.01208177674561739
step: 130, loss: 0.011535262688994408
step: 140, loss: 0.01761784590780735
step: 150, loss: 0.004168766550719738
step: 160, loss: 0.007499401923269033
step: 170, loss: 0.007282689213752747
step: 180, loss: 0.08235646039247513
step: 190, loss: 0.00559643330052495
step: 200, loss: 0.012778094038367271
step: 210, loss: 0.002917024539783597
/usr/local/lib/python3.7/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=0.5210084033613446, f1=0.556745182012848,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006598450243473053
step: 10, loss: 0.006863533519208431
step: 20, loss: 0.016214944422245026
step: 30, loss: 0.07595513761043549
step: 40, loss: 0.006596136372536421
step: 50, loss: 0.013546541333198547
step: 60, loss: 0.02600296027958393
step: 70, loss: 0.091685950756073
step: 80, loss: 0.08521749079227448
step: 90, loss: 0.1087002083659172
step: 100, loss: 0.015337453223764896
step: 110, loss: 0.09749875962734222
step: 120, loss: 0.011073780246078968
step: 130, loss: 0.0018517550779506564
step: 140, loss: 0.0035800253972411156
step: 150, loss: 0.017145724967122078
step: 160, loss: 0.023487208411097527
step: 170, loss: 0.06866085529327393
step: 180, loss: 0.003015568945556879
step: 190, loss: 0.005515769589692354
step: 200, loss: 0.0066055660136044025
step: 210, loss: 0.020048122853040695
/usr/local/lib/python3.7/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=0.5247933884297521, f1=0.5583333333333333,
best_f1=0.5525114155251142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020674618426710367
step: 10, loss: 0.02343863621354103
step: 20, loss: 0.0011555776000022888
step: 30, loss: 0.0044577536173164845
step: 40, loss: 0.0010338168358430266
step: 50, loss: 0.006878018379211426
step: 60, loss: 0.08681309223175049
step: 70, loss: 0.010602572932839394
step: 80, loss: 0.03541447222328186
step: 90, loss: 0.01402830146253109
step: 100, loss: 0.003958313725888729
step: 110, loss: 0.001737932674586773
step: 120, loss: 0.03465226665139198
step: 130, loss: 0.08717940747737885
step: 140, loss: 0.006398422177881002
step: 150, loss: 0.024936974048614502
step: 160, loss: 0.012522895820438862
step: 170, loss: 0.08613845705986023
step: 180, loss: 0.014675255864858627
step: 190, loss: 0.03317979350686073
step: 200, loss: 0.03336094692349434
step: 210, loss: 0.05189982056617737
/usr/local/lib/python3.7/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=0.5243128964059196, f1=0.546236559139785,
best_f1=0.5525114155251142
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 [00:03, 617.95it/s]
load_f1 = 0.5870445344129555
real_f1 = 0.579476861167002
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 [00:10, 401.30it/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 distilbert \
--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 distilbert \  
--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: 0.8583292961120605
```

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

```
step: 10, loss: 0.16320469975471497
```

```
step: 20, loss: 0.15632113814353943
```

```
step: 30, loss: 0.5212987065315247
```

```

step: 40, loss: 0.2688625454902649
step: 50, loss: 0.30246198177337646
step: 60, loss: 0.36781755089759827
step: 70, loss: 0.18569298088550568
step: 80, loss: 0.5423967242240906
step: 90, loss: 0.2390148788690567
step: 100, loss: 0.22887130081653595
step: 110, loss: 0.2341650426387787
step: 120, loss: 0.41608473658561707
step: 130, loss: 0.34886497259140015
step: 140, loss: 0.3376144766807556
step: 150, loss: 0.27797549962997437
step: 160, loss: 0.21412082016468048
step: 170, loss: 0.402037650346756
step: 180, loss: 0.3056505620479584
step: 190, loss: 0.11429216712713242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.44282238442822386, f1=0.48648648648648646,
best_f1=0.48648648648648646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3205488920211792
step: 10, loss: 0.05770498514175415
step: 20, loss: 0.09020356088876724
step: 30, loss: 0.23598612844944
step: 40, loss: 0.4641963541507721
step: 50, loss: 0.3615866005420685
step: 60, loss: 0.13854742050170898
step: 70, loss: 0.35478177666664124
step: 80, loss: 0.1737249493598938
step: 90, loss: 0.10449942201375961
step: 100, loss: 0.30909085273742676
step: 110, loss: 0.11691775918006897
step: 120, loss: 0.2064087837934494
step: 130, loss: 0.1465838998556137
step: 140, loss: 0.21844270825386047
step: 150, loss: 0.05219445005059242
step: 160, loss: 0.12340570986270905
step: 170, loss: 0.2123776227235794
step: 180, loss: 0.21838384866714478
step: 190, loss: 0.1538853645324707

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7055555555555556, f1=0.7029972752043598,
best_f1=0.7029972752043598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12633146345615387
step: 10, loss: 0.19126452505588531
step: 20, loss: 0.07063858956098557
step: 30, loss: 0.036269411444664
step: 40, loss: 0.05056171119213104
step: 50, loss: 0.13844162225723267
step: 60, loss: 0.05798553302884102
step: 70, loss: 0.2525928318500519
step: 80, loss: 0.18990077078342438
step: 90, loss: 0.035932641476392746
step: 100, loss: 0.09712245315313339
step: 110, loss: 0.28600287437438965
step: 120, loss: 0.04669606685638428
step: 130, loss: 0.052258022129535675
step: 140, loss: 0.08668265491724014
step: 150, loss: 0.1449359506368637
step: 160, loss: 0.07922901958227158
step: 170, loss: 0.06663965433835983
step: 180, loss: 0.1310385763645172
step: 190, loss: 0.18737904727458954
/usr/local/lib/python3.7/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=0.705539358600583, f1=0.6707692307692308,
best_f1=0.7029972752043598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04847058281302452
step: 10, loss: 0.07809042930603027
step: 20, loss: 0.011599624529480934
step: 30, loss: 0.07595908641815186
step: 40, loss: 0.008978270925581455
step: 50, loss: 0.0846613422036171
step: 60, loss: 0.1691371500492096
step: 70, loss: 0.00930067989975214

```



```

step: 80, loss: 0.05862070992588997
step: 90, loss: 0.027326365932822227
step: 100, loss: 0.06504945456981659
step: 110, loss: 0.013706411235034466
step: 120, loss: 0.03448636829853058
step: 130, loss: 0.1391776204109192
step: 140, loss: 0.05945092812180519
step: 150, loss: 0.02899528667330742
step: 160, loss: 0.05005710572004318
step: 170, loss: 0.08389075100421906
step: 180, loss: 0.04206841066479683
step: 190, loss: 0.03769475966691971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.744186046511628, f1=0.7437185929648241,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21261988580226898
step: 10, loss: 0.02395492233335972
step: 20, loss: 0.13430339097976685
step: 30, loss: 0.025939829647541046
step: 40, loss: 0.1332826018333435
step: 50, loss: 0.024941110983490944
step: 60, loss: 0.0292144063860178
step: 70, loss: 0.024751784279942513
step: 80, loss: 0.031550657004117966
step: 90, loss: 0.0037563610821962357
step: 100, loss: 0.09896247088909149
step: 110, loss: 0.02227463386952877
step: 120, loss: 0.006007807794958353
step: 130, loss: 0.021828526630997658
step: 140, loss: 0.014357135631144047
step: 150, loss: 0.023204462602734566
step: 160, loss: 0.00444471649825573
step: 170, loss: 0.0873127430677414
step: 180, loss: 0.21426890790462494
step: 190, loss: 0.1723581850528717
/usr/local/lib/python3.7/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=0.7073791348600508, f1=0.7393617021276595,

```

```

best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028419531881809235
step: 10, loss: 0.002232207916676998
step: 20, loss: 0.019585587084293365
step: 30, loss: 0.008823037147521973
step: 40, loss: 0.07197390496730804
step: 50, loss: 0.029844798147678375
step: 60, loss: 0.005606155842542648
step: 70, loss: 0.04631977155804634
step: 80, loss: 0.026373019441962242
step: 90, loss: 0.00883230846375227
step: 100, loss: 0.003256440395489335
step: 110, loss: 0.03715882822871208
step: 120, loss: 0.0600966140627861
step: 130, loss: 0.049354083836078644
step: 140, loss: 0.0240352563560009
step: 150, loss: 0.02144618332386017
step: 160, loss: 0.0032682158052921295
step: 170, loss: 0.017859667539596558
step: 180, loss: 0.004453715868294239
step: 190, loss: 0.10602477937936783
/usr/local/lib/python3.7/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=0.7277936962750716, f1=0.7277936962750716,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014512666501104832
step: 10, loss: 0.10342574864625931
step: 20, loss: 0.01051039807498455
step: 30, loss: 0.009780018590390682
step: 40, loss: 0.0170400682836771
step: 50, loss: 0.0273960642516613
step: 60, loss: 0.016631266102194786
step: 70, loss: 0.0029617163818329573
step: 80, loss: 0.1461264044046402
step: 90, loss: 0.0059165493585169315
step: 100, loss: 0.01196776982396841
step: 110, loss: 0.00630069337785244
step: 120, loss: 0.013853165321052074
step: 130, loss: 0.056765735149383545
step: 140, loss: 0.010804006829857826
step: 150, loss: 0.02454224042594433

```

```

step: 160, loss: 0.007458561100065708
step: 170, loss: 0.07686929404735565
step: 180, loss: 0.008606350049376488
step: 190, loss: 0.011291619390249252
/usr/local/lib/python3.7/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=0.7351351351351353, f1=0.7486338797814207,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007560830097645521
step: 10, loss: 0.005570015404373407
step: 20, loss: 0.03034953773021698
step: 30, loss: 0.002242974704131484
step: 40, loss: 0.049719344824552536
step: 50, loss: 0.0014583439333364367
step: 60, loss: 0.0004497472837101668
step: 70, loss: 0.010436800308525562
step: 80, loss: 0.1525442898273468
step: 90, loss: 0.0004925057874061167
step: 100, loss: 0.027332527562975883
step: 110, loss: 0.0021629375405609608
step: 120, loss: 0.007689088117331266
step: 130, loss: 0.004362210631370544
step: 140, loss: 0.04533912613987923
step: 150, loss: 0.0029366211965680122
step: 160, loss: 0.003524979343637824
step: 170, loss: 0.0030199552420526743
step: 180, loss: 0.02500610053539276
step: 190, loss: 0.009990940801799297
/usr/local/lib/python3.7/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=0.7182320441988949, f1=0.7267605633802817,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034667328000068665
step: 10, loss: 0.0016326976474374533
step: 20, loss: 0.01792120188474655
step: 30, loss: 0.012873937375843525
step: 40, loss: 0.0016242306446656585
step: 50, loss: 0.0013314560055732727
step: 60, loss: 0.004190295469015837
step: 70, loss: 0.04261571541428566

```

```

step: 80, loss: 0.018654031679034233
step: 90, loss: 0.04274386167526245
step: 100, loss: 0.003657687222585082
step: 110, loss: 0.004015607293695211
step: 120, loss: 0.01939929835498333
step: 130, loss: 0.0025649918243288994
step: 140, loss: 0.0015947960782796144
step: 150, loss: 0.04341813921928406
step: 160, loss: 0.0010008576791733503
step: 170, loss: 0.003306680591776967
step: 180, loss: 0.01611752063035965
step: 190, loss: 0.0022833910770714283
/usr/local/lib/python3.7/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=0.7098591549295775, f1=0.7308781869688386,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000542731024324894
step: 10, loss: 0.019077176228165627
step: 20, loss: 0.0006533603882417083
step: 30, loss: 0.007166887633502483
step: 40, loss: 0.00047399912727996707
step: 50, loss: 0.0974709764122963
step: 60, loss: 0.03638928756117821
step: 70, loss: 0.0013838913291692734
step: 80, loss: 0.0006320816464722157
step: 90, loss: 0.0040892865508794785
step: 100, loss: 0.007366596255451441
step: 110, loss: 0.012646768242120743
step: 120, loss: 0.028054846450686455
step: 130, loss: 0.0022404545452445745
step: 140, loss: 0.030699661001563072
step: 150, loss: 0.0005291855777613819
step: 160, loss: 0.0003029990184586495
step: 170, loss: 0.015809910371899605
step: 180, loss: 0.20481882989406586
step: 190, loss: 0.002738049952313304
/usr/local/lib/python3.7/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=0.7231638418079097, f1=0.7492957746478874,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0009873652597889304
step: 10, loss: 0.0005698500317521393
step: 20, loss: 0.07860992103815079
step: 30, loss: 0.03424472361803055
step: 40, loss: 0.0016142262611538172
step: 50, loss: 0.0012263195822015405
step: 60, loss: 0.0011208751238882542
step: 70, loss: 0.000383375067031011
step: 80, loss: 0.0007248633191920817
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.005753586068749428
step: 100, loss: 0.0034417794086039066
step: 110, loss: 0.003924371208995581
step: 120, loss: 0.004631347954273224
step: 130, loss: 0.0023054350167512894
step: 140, loss: 0.0017440839437767863
step: 150, loss: 0.0015328193549066782
step: 160, loss: 0.000805037678219378
step: 170, loss: 0.00770717766135931
step: 180, loss: 0.019793665036559105
step: 190, loss: 0.0039808242581784725
/usr/local/lib/python3.7/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=0.7277628032345013, f1=0.7377049180327869,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002002178458496928
step: 10, loss: 0.0027519885916262865
step: 20, loss: 0.021313967183232307
step: 30, loss: 0.0013820872409269214
step: 40, loss: 0.010652505792677402
step: 50, loss: 0.0024754221085458994
step: 60, loss: 0.002899318002164364
step: 70, loss: 0.0021728852298110723
step: 80, loss: 0.0007973472820594907
step: 90, loss: 0.03100700117647648
step: 100, loss: 0.0016963421367108822
step: 110, loss: 0.0009051532833836973
step: 120, loss: 0.0007732806843705475
step: 130, loss: 0.001832361682318151
step: 140, loss: 0.004838001914322376
step: 150, loss: 0.002299506915733218
step: 160, loss: 0.015203610993921757
step: 170, loss: 0.07366202771663666
step: 180, loss: 0.0008151550428010523

```

```

step: 190, loss: 0.016430819407105446
/usr/local/lib/python3.7/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=0.7247191011235956, f1=0.7394957983193278,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007920459029264748
step: 10, loss: 0.13708332180976868
step: 20, loss: 0.0021577272564172745
step: 30, loss: 0.0009051139350049198
step: 40, loss: 0.0002693583373911679
step: 50, loss: 0.0013715183595195413
step: 60, loss: 0.001165095018222928
step: 70, loss: 0.0009094592533074319
step: 80, loss: 0.12018583714962006
step: 90, loss: 0.004045750014483929
step: 100, loss: 0.003947989083826542
step: 110, loss: 0.0017419657669961452
step: 120, loss: 0.005215339362621307
step: 130, loss: 0.006809697952121496
step: 140, loss: 0.010431302711367607
step: 150, loss: 0.008411196991801262
step: 160, loss: 0.0021947838831692934
step: 170, loss: 0.00019986323604825884
step: 180, loss: 0.02495443820953369
step: 190, loss: 0.004024272318929434
/usr/local/lib/python3.7/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=0.7232876712328767, f1=0.7403314917127072,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.044804394245147705
step: 10, loss: 0.00050976756028831
step: 20, loss: 0.004075200296938419
step: 30, loss: 0.0017798744374886155
step: 40, loss: 0.0003276584902778268
step: 50, loss: 0.0011238149600103498
step: 60, loss: 0.002306931419298053
step: 70, loss: 0.02648773230612278
step: 80, loss: 0.0003676906053442508
step: 90, loss: 0.007986842654645443
step: 100, loss: 0.008514802902936935

```

```

step: 110, loss: 0.0008408937137573957
step: 120, loss: 0.05440900847315788
step: 130, loss: 0.0004609638708643615
step: 140, loss: 0.0003306026046629995
step: 150, loss: 0.0003683150280267
step: 160, loss: 0.00036679196637123823
step: 170, loss: 0.004632245283573866
step: 180, loss: 0.0019956857431679964
step: 190, loss: 0.0008605080656707287
/usr/local/lib/python3.7/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=0.7331536388140162, f1=0.745945945945946,
best_f1=0.7437185929648241
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014129397459328175
step: 10, loss: 0.00039234725409187376
step: 20, loss: 0.06945198029279709
step: 30, loss: 0.0019384329207241535
step: 40, loss: 0.04187460243701935
step: 50, loss: 0.004364873748272657
step: 60, loss: 0.0015430077910423279
step: 70, loss: 0.0005045346333645284
step: 80, loss: 0.0008614075486548245
step: 90, loss: 0.0018457736587151885
step: 100, loss: 0.0004982258542440832
step: 110, loss: 0.00175397377461195
step: 120, loss: 0.0007466251263394952
step: 130, loss: 0.0008755954913794994
step: 140, loss: 0.1597178727388382
step: 150, loss: 0.0013057386968284845
step: 160, loss: 0.02045348472893238
step: 170, loss: 0.0009279417572543025
step: 180, loss: 0.0017173400847241282
step: 190, loss: 0.015329363755881786
/usr/local/lib/python3.7/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=0.7303370786516852, f1=0.7450980392156863,
best_f1=0.7437185929648241
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 [00:06, 337.00it/s]
load_f1 = 0.6201923076923077
real_f1 = 0.6132075471698113
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 [00:11, 394.59it/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 distilbert \
--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 distilbert \
  --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: 0.856614887714386

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

step: 10, loss: 0.23099952936172485

step: 20, loss: 0.1537337303161621

step: 30, loss: 0.235469788312912

step: 40, loss: 0.31415557861328125

step: 50, loss: 0.38184016942977905

step: 60, loss: 0.45045480132102966

step: 70, loss: 0.31452351808547974

step: 80, loss: 0.24690210819244385

step: 90, loss: 0.40372732281684875

step: 100, loss: 0.23741847276687622

step: 110, loss: 0.18858343362808228

step: 120, loss: 0.54661625623703

step: 130, loss: 0.41729456186294556

step: 140, loss: 0.46534836292266846

step: 150, loss: 0.09596347063779831

step: 160, loss: 0.2573399841785431

step: 170, loss: 0.1828957200050354

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

An input tensor was not cuda.

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

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

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=0.6326034063260341, f1=0.6496519721577726,

best_f1=0.6496519721577726

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

An input tensor was not cuda.

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

step: 0, loss: 0.24427597224712372

step: 10, loss: 0.06738138943910599

step: 20, loss: 0.24622061848640442

step: 30, loss: 0.09769641607999802

step: 40, loss: 0.15235549211502075

step: 50, loss: 0.18264830112457275

step: 60, loss: 0.09036286920309067

```

step: 70, loss: 0.1326739639043808
step: 80, loss: 0.09591496735811234
step: 90, loss: 0.28702130913734436
step: 100, loss: 0.08382274210453033
step: 110, loss: 0.20614077150821686
step: 120, loss: 0.10178809612989426
step: 130, loss: 0.12654784321784973
step: 140, loss: 0.1978316456079483
step: 150, loss: 0.11985749751329422
step: 160, loss: 0.0864110216498375
step: 170, loss: 0.07788300514221191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7469879518072288, f1=0.7467811158798283,
best_f1=0.7467811158798283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1217934787273407
step: 10, loss: 0.08198040723800659
step: 20, loss: 0.011268585920333862
step: 30, loss: 0.15513485670089722
step: 40, loss: 0.012414058670401573
step: 50, loss: 0.08178305625915527
step: 60, loss: 0.16230201721191406
step: 70, loss: 0.07234040647745132
step: 80, loss: 0.28340327739715576
step: 90, loss: 0.07563565671443939
step: 100, loss: 0.08126115053892136
step: 110, loss: 0.0789579227566719
step: 120, loss: 0.010257597081363201
step: 130, loss: 0.14037778973579407
step: 140, loss: 0.005291856825351715
step: 150, loss: 0.04443485662341118
step: 160, loss: 0.03288819640874863
step: 170, loss: 0.08233121037483215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7546174142480211, f1=0.7785888077858881,
best_f1=0.7785888077858881
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11419542133808136
step: 10, loss: 0.19237209856510162
step: 20, loss: 0.06675665825605392
step: 30, loss: 0.03871093690395355
step: 40, loss: 0.006677665747702122
step: 50, loss: 0.02910703420639038
step: 60, loss: 0.15771709382534027
step: 70, loss: 0.011133810505270958
step: 80, loss: 0.054710403084754944
step: 90, loss: 0.08778833597898483
step: 100, loss: 0.019916966557502747
step: 110, loss: 0.019321786239743233
step: 120, loss: 0.07338784635066986
step: 130, loss: 0.07593680918216705
step: 140, loss: 0.023288076743483543
step: 150, loss: 0.003326146863400936
step: 160, loss: 0.07089883834123611
step: 170, loss: 0.008620526641607285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7821782178217821, f1=0.786206896551724,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025963786989450455
step: 10, loss: 0.03388659656047821
step: 20, loss: 0.028409631922841072
step: 30, loss: 0.06332173198461533
step: 40, loss: 0.031061258167028427
step: 50, loss: 0.014675614424049854
step: 60, loss: 0.0009915076661854982
step: 70, loss: 0.13351276516914368
step: 80, loss: 0.012636450119316578
step: 90, loss: 0.027131495997309685
step: 100, loss: 0.02042163908481598
step: 110, loss: 0.2985367178916931
step: 120, loss: 0.01222661230713129

```

```

step: 130, loss: 0.0041253454983234406
step: 140, loss: 0.022892428562045097
step: 150, loss: 0.0063021001406013966
step: 160, loss: 0.17743968963623047
step: 170, loss: 0.14961011707782745
/usr/local/lib/python3.7/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=0.7479674796747967, f1=0.745679012345679,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0492277555167675
step: 10, loss: 0.07035721838474274
step: 20, loss: 0.006367420312017202
step: 30, loss: 0.03330202028155327
step: 40, loss: 0.0029126086737960577
step: 50, loss: 0.014476332813501358
step: 60, loss: 0.13355709612369537
step: 70, loss: 0.006294353399425745
step: 80, loss: 0.15881463885307312
step: 90, loss: 0.051922041922807693
step: 100, loss: 0.09070105105638504
step: 110, loss: 0.047228842973709106
step: 120, loss: 0.03311098366975784
step: 130, loss: 0.20123524963855743
step: 140, loss: 0.003358403919264674
step: 150, loss: 0.18431490659713745
step: 160, loss: 0.06711885333061218
step: 170, loss: 0.012129496783018112
/usr/local/lib/python3.7/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=0.760705289672544, f1=0.7788018433179723,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.051835689693689346
step: 10, loss: 0.0023943977430462837
step: 20, loss: 0.011040025390684605
step: 30, loss: 0.0023168621119111776
step: 40, loss: 0.016864357516169548
step: 50, loss: 0.01673927716910839
step: 60, loss: 0.2751353085041046
step: 70, loss: 0.02450498752295971
step: 80, loss: 0.002703897189348936

```

```

step: 90, loss: 0.00916727539151907
step: 100, loss: 0.0023776055313646793
step: 110, loss: 0.009563468396663666
step: 120, loss: 0.19957894086837769
step: 130, loss: 0.20443306863307953
step: 140, loss: 0.03866996988654137
step: 150, loss: 0.10958194732666016
step: 160, loss: 0.03401133045554161
step: 170, loss: 0.06397856771945953
/usr/local/lib/python3.7/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=0.7494033412887827, f1=0.7405764966740576,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018746377900242805
step: 10, loss: 0.011236493475735188
step: 20, loss: 0.006346737965941429
step: 30, loss: 0.031115315854549408
step: 40, loss: 0.003766231471672654
step: 50, loss: 0.002015621867030859
step: 60, loss: 0.023731263354420662
step: 70, loss: 0.04082757979631424
step: 80, loss: 0.0061518424190580845
step: 90, loss: 0.10972137004137039
step: 100, loss: 0.010378310456871986
step: 110, loss: 0.0772729218006134
step: 120, loss: 0.006326995324343443
step: 130, loss: 0.004923928529024124
step: 140, loss: 0.0007436537416651845
step: 150, loss: 0.06260013580322266
step: 160, loss: 0.044406622648239136
step: 170, loss: 0.001141666783951223
/usr/local/lib/python3.7/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=0.7628865979381445, f1=0.7813953488372092,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01960724964737892
step: 10, loss: 0.006778951268643141
step: 20, loss: 0.007931227795779705
step: 30, loss: 0.008254104293882847
step: 40, loss: 0.010286071337759495

```

```

step: 50, loss: 0.015535622835159302
step: 60, loss: 0.0010520063806325197
step: 70, loss: 0.001235035597346723
step: 80, loss: 0.01955391839146614
step: 90, loss: 0.02670985832810402
step: 100, loss: 0.019605794921517372
step: 110, loss: 0.013786916621029377
step: 120, loss: 0.010270926170051098
step: 130, loss: 0.00016415216668974608
step: 140, loss: 0.00042920553823933005
step: 150, loss: 0.006968047469854355
step: 160, loss: 0.015182333067059517
step: 170, loss: 0.056245021522045135
/usr/local/lib/python3.7/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=0.7499999999999999, f1=0.7703016241299303,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00405809236690402
step: 10, loss: 0.015745893120765686
step: 20, loss: 0.008397595956921577
step: 30, loss: 0.0006994510185904801
step: 40, loss: 0.13156330585479736
step: 50, loss: 0.053108636289834976
step: 60, loss: 0.001968545373529196
step: 70, loss: 0.002016617450863123
step: 80, loss: 0.010630532167851925
step: 90, loss: 0.10331973433494568
step: 100, loss: 0.09419557452201843
step: 110, loss: 0.013097111135721207
step: 120, loss: 0.021576525643467903
step: 130, loss: 0.01169810350984335
step: 140, loss: 0.01686665415763855
step: 150, loss: 0.002134501002728939
step: 160, loss: 0.00110304553527385
step: 170, loss: 0.0012458594283089042
/usr/local/lib/python3.7/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=0.7777777777777778, f1=0.7731481481481483,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006104345084168017

```

```

step: 10, loss: 0.08904781937599182
step: 20, loss: 0.028497202321887016
step: 30, loss: 0.008336210623383522
step: 40, loss: 0.010646730661392212
step: 50, loss: 0.009128994308412075
step: 60, loss: 0.0032184000592678785
step: 70, loss: 0.008005263283848763
step: 80, loss: 0.0009621887002140284
step: 90, loss: 0.0017094825161620975
step: 100, loss: 0.0002504177391529083
step: 110, loss: 0.006183742079883814
step: 120, loss: 0.0038491678424179554
step: 130, loss: 0.00257997983135283
step: 140, loss: 0.06417754292488098
step: 150, loss: 0.03260368853807449
step: 160, loss: 0.0028414526022970676
step: 170, loss: 0.07759380340576172
/usr/local/lib/python3.7/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=0.7589743589743588, f1=0.7688679245283018,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003746790112927556
step: 10, loss: 0.011133203282952309
step: 20, loss: 0.025279739871621132
step: 30, loss: 0.00137422070838511
step: 40, loss: 0.002178823808208108
step: 50, loss: 0.0006356448866426945
step: 60, loss: 0.004247057251632214
step: 70, loss: 0.010268889367580414
step: 80, loss: 0.027381494641304016
step: 90, loss: 0.0014901814283803105
step: 100, loss: 0.0041030943393707275
step: 110, loss: 0.00037664928822778165
step: 120, loss: 0.0437777079641819
step: 130, loss: 0.004811463877558708
step: 140, loss: 0.00035513684269972146
step: 150, loss: 0.04572010040283203
step: 160, loss: 0.00378427398391068
step: 170, loss: 0.0011203242465853691
/usr/local/lib/python3.7/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=0.7777777777777778, f1=0.7551487414187643,
best_f1=0.786206896551724

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004203777643851936
step: 10, loss: 0.03919994458556175
step: 20, loss: 0.0033920882269740105
step: 30, loss: 0.00034089115797542036
step: 40, loss: 0.00020765623776242137
step: 50, loss: 0.009743678383529186
step: 60, loss: 0.0005463010165840387
step: 70, loss: 0.1322222650051117
step: 80, loss: 0.06521115452051163
step: 90, loss: 0.0004521296941675246
step: 100, loss: 0.0026470297016203403
step: 110, loss: 0.0028687510639429092
step: 120, loss: 0.0012806684244424105
step: 130, loss: 0.0022648724261671305
step: 140, loss: 0.00029112541233189404
step: 150, loss: 0.006522546522319317
step: 160, loss: 0.0016277176328003407
step: 170, loss: 0.004854250233620405
/usr/local/lib/python3.7/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=0.7688564476885645, f1=0.7324561403508772,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016286917962133884
step: 10, loss: 0.0006925841444171965
step: 20, loss: 0.05136801302433014
step: 30, loss: 0.001344584976322949
step: 40, loss: 0.00042220274917781353
step: 50, loss: 0.00904912780970335
step: 60, loss: 0.0003175001766066998
step: 70, loss: 0.0055464389733970165
step: 80, loss: 0.0007651323103345931
step: 90, loss: 0.0008043188718147576
step: 100, loss: 0.0007177574443630874
step: 110, loss: 0.0008937451057136059
step: 120, loss: 0.018346788361668587
step: 130, loss: 0.0008304562652483582
step: 140, loss: 0.027399715036153793
step: 150, loss: 0.006439403165131807
step: 160, loss: 0.03578481823205948
step: 170, loss: 0.0009075817069970071
/usr/local/lib/python3.7/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=0.766839378238342, f1=0.7663551401869159,
best_f1=0.786206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032926187850534916
step: 10, loss: 0.002266670111566782
step: 20, loss: 0.01616467908024788
step: 30, loss: 0.02876443788409233
step: 40, loss: 0.0003575480659492314
step: 50, loss: 0.004601619206368923
step: 60, loss: 0.0027628098614513874
step: 70, loss: 0.011605649255216122
step: 80, loss: 0.0007610134198330343
step: 90, loss: 0.00041926943231374025
step: 100, loss: 0.0019443457713350654
step: 110, loss: 0.0004480917123146355
step: 120, loss: 0.09437574446201324
step: 130, loss: 0.0018029737984761596
step: 140, loss: 0.0003991895937360823
step: 150, loss: 0.015054458752274513
step: 160, loss: 0.0008857196662575006
step: 170, loss: 0.0014902843395248055
/usr/local/lib/python3.7/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=0.7692307692307693, f1=0.7575057736720553,
best_f1=0.786206896551724
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 [00:04, 438.68it/s]
load_f1 = 0.3801065719360568
real_f1 = 0.33684210526315794
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 [00:11, 391.89it/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 distilbert \
    --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 distilbert \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Downloading: 100% 442/442 [00:00<00:00, 413kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 1.22MB/s]
 Downloading: 100% 268M/268M [00:15<00:00, 17.8MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```
enabled          : True
opt_level        : 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: 0.8197415471076965
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4796507954597473
step: 20, loss: 0.5769311189651489
step: 30, loss: 0.4711677134037018
step: 40, loss: 0.31284284591674805
step: 50, loss: 0.19791676104068756
step: 60, loss: 0.12296945601701736
step: 70, loss: 0.16037636995315552
step: 80, loss: 0.19414538145065308
step: 90, loss: 0.06059332937002182
step: 100, loss: 0.12323015183210373
step: 110, loss: 0.12344454973936081
step: 120, loss: 0.02401423454284668
step: 130, loss: 0.03749920055270195
step: 140, loss: 0.03123793937265873
step: 150, loss: 0.23427970707416534
step: 160, loss: 0.28940558433532715
step: 170, loss: 0.032351452857255936
step: 180, loss: 0.003019610419869423
step: 190, loss: 0.08718789368867874
step: 200, loss: 0.012913295067846775
step: 210, loss: 0.02509474940598011
step: 220, loss: 0.012131858617067337
step: 230, loss: 0.02609637752175331
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9502762430939226, f1=0.9433962264150944,
best_f1=0.9433962264150944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0412575788795948
step: 10, loss: 0.007646108046174049
step: 20, loss: 0.03800078481435776
step: 30, loss: 0.004776206333190203
step: 40, loss: 0.03307240083813667
step: 50, loss: 0.038154806941747665
step: 60, loss: 0.016188226640224457
step: 70, loss: 0.03690938651561737
step: 80, loss: 0.0022441092878580093
step: 90, loss: 0.007845258340239525

```

```

step: 100, loss: 0.12879571318626404
step: 110, loss: 0.13277091085910797
step: 120, loss: 0.10684884339570999
step: 130, loss: 0.025694172829389572
step: 140, loss: 0.07963715493679047
step: 150, loss: 0.04414456710219383
step: 160, loss: 0.008523093536496162
step: 170, loss: 0.11066891252994537
step: 180, loss: 0.015170471742749214
step: 190, loss: 0.15980055928230286
step: 200, loss: 0.022417914122343063
step: 210, loss: 0.057252734899520874
step: 220, loss: 0.003440933534875512
step: 230, loss: 0.03134509176015854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9636163175303197, f1=0.9556541019955654,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07131925225257874
step: 10, loss: 0.011024809442460537
step: 20, loss: 0.0029610004276037216
step: 30, loss: 0.05713942274451256
step: 40, loss: 0.2407711148262024
step: 50, loss: 0.008970011956989765
step: 60, loss: 0.03953753039240837
step: 70, loss: 0.0016598141519352794
step: 80, loss: 0.0013727708719670773
step: 90, loss: 0.1378149539232254
step: 100, loss: 0.025450050830841064
step: 110, loss: 0.01775587722659111
step: 120, loss: 0.0062019191682338715
step: 130, loss: 0.0004922968801110983
step: 140, loss: 0.013072509318590164
step: 150, loss: 0.0028037927113473415
step: 160, loss: 0.017628556117415428
step: 170, loss: 0.007110767997801304
step: 180, loss: 0.013221598230302334
step: 190, loss: 0.0031866144854575396
step: 200, loss: 0.14761506021022797
step: 210, loss: 0.005218996200710535

```

```

step: 220, loss: 0.00862483587116003
step: 230, loss: 0.009116780944168568
/usr/local/lib/python3.7/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=0.9620535714285715, f1=0.9605411499436302,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004417491611093283
step: 10, loss: 0.02087957039475441
step: 20, loss: 0.0040977513417601585
step: 30, loss: 0.001397728337906301
step: 40, loss: 0.022036878392100334
step: 50, loss: 0.02641647309064865
step: 60, loss: 0.0015779630048200488
step: 70, loss: 0.005959283094853163
step: 80, loss: 0.14208440482616425
step: 90, loss: 0.00881465058773756
step: 100, loss: 0.00533232931047678
step: 110, loss: 0.022900845855474472
step: 120, loss: 0.0036264020018279552
step: 130, loss: 0.004255683161318302
step: 140, loss: 0.01655227690935135
step: 150, loss: 0.0018833904759958386
step: 160, loss: 0.0037948135286569595
step: 170, loss: 0.015623507089912891
step: 180, loss: 0.13780106604099274
step: 190, loss: 0.009731859900057316
step: 200, loss: 0.0009194401209242642
step: 210, loss: 0.0008754490991123021
step: 220, loss: 0.0004960219375789165
step: 230, loss: 0.0004191266489215195
/usr/local/lib/python3.7/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=0.9634551495016611, f1=0.9556541019955654,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012236346490681171
step: 10, loss: 0.006407981272786856
step: 20, loss: 0.0007129760924726725
step: 30, loss: 0.0004678376717492938
step: 40, loss: 0.0046886601485311985
step: 50, loss: 0.0005718168104067445

```

```

step: 60, loss: 0.0014433461474254727
step: 70, loss: 0.0006156049203127623
step: 80, loss: 0.0008352843578904867
step: 90, loss: 0.0003571966663002968
step: 100, loss: 0.0019846188370138407
step: 110, loss: 0.004559106193482876
step: 120, loss: 0.04373747855424881
step: 130, loss: 0.12588436901569366
step: 140, loss: 0.008257579989731312
step: 150, loss: 0.00031307421158999205
step: 160, loss: 0.015028099529445171
step: 170, loss: 0.10835115611553192
step: 180, loss: 0.0037248432636260986
step: 190, loss: 0.00511428527534008
step: 200, loss: 0.01681843213737011
step: 210, loss: 0.0026987630408257246
step: 220, loss: 0.0033477921970188618
step: 230, loss: 0.0031775753013789654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9655172413793103, f1=0.9611542730299667,
best_f1=0.9611542730299667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011121598072350025
step: 10, loss: 0.04410089924931526
step: 20, loss: 0.004225470125675201
step: 30, loss: 0.0004904899396933615
step: 40, loss: 0.11840817332267761
step: 50, loss: 0.00400158204138279
step: 60, loss: 0.08455432206392288
step: 70, loss: 0.0008927752496674657
step: 80, loss: 0.0037028116639703512
step: 90, loss: 0.0022542919032275677
step: 100, loss: 0.022012358531355858
step: 110, loss: 0.027148321270942688
step: 120, loss: 0.0007947620470076799
step: 130, loss: 0.00029589043697342277
step: 140, loss: 0.0016803056932985783
step: 150, loss: 0.0003414525417611003
step: 160, loss: 0.0011822838569059968
step: 170, loss: 0.00015673924644943327

```



```

step: 180, loss: 0.0003748904273379594
step: 190, loss: 0.0005876322975382209
step: 200, loss: 0.1310313194990158
step: 210, loss: 0.0022976319305598736
step: 220, loss: 0.0007781014428474009
step: 230, loss: 0.0006214501336216927
/usr/local/lib/python3.7/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=0.9650507328072153, f1=0.9627959413754228,
best_f1=0.9611542730299667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03446407988667488
step: 10, loss: 0.0009704334079287946
step: 20, loss: 0.0014265954960137606
step: 30, loss: 0.0023516137152910233
step: 40, loss: 0.02280411124229431
step: 50, loss: 0.00044988704030402005
step: 60, loss: 0.003459582570940256
step: 70, loss: 0.000505134928971529
step: 80, loss: 0.014198661781847477
step: 90, loss: 0.0007112532039172947
step: 100, loss: 0.0013846437213942409
step: 110, loss: 0.0020013817120343447
step: 120, loss: 0.0013438722817227244
step: 130, loss: 0.001664430252276361
step: 140, loss: 0.00015497342974413186
step: 150, loss: 0.006524214521050453
step: 160, loss: 0.001596699352376163
step: 170, loss: 0.00030631248955614865
step: 180, loss: 0.000985517748631537
step: 190, loss: 0.004063686355948448
step: 200, loss: 0.01180701982229948
step: 210, loss: 0.0017003766261041164
step: 220, loss: 0.0011397070484235883
step: 230, loss: 0.005299909971654415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9664429530201343, f1=0.9597315436241611,
best_f1=0.9597315436241611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019072556868195534
step: 10, loss: 0.023266222327947617
step: 20, loss: 0.00033408342278562486
step: 30, loss: 0.00027135186246596277
step: 40, loss: 0.0009389488259330392
step: 50, loss: 0.0007753908867016435
step: 60, loss: 0.0002619594451971352
step: 70, loss: 0.016670580953359604
step: 80, loss: 0.0009381707641296089
step: 90, loss: 0.0006198559422045946
step: 100, loss: 0.00016055202286224812
step: 110, loss: 0.00014224392361938953
step: 120, loss: 7.951282896101475e-05
step: 130, loss: 0.0003101903130300343
step: 140, loss: 0.010079389438033104
step: 150, loss: 0.0002831669698935002
step: 160, loss: 0.009040656499564648
step: 170, loss: 0.0012052454985678196
step: 180, loss: 0.00032774455030448735
step: 190, loss: 0.0014237447176128626
step: 200, loss: 0.051550112664699554
step: 210, loss: 0.00030492566293105483
step: 220, loss: 0.0011699895840138197
step: 230, loss: 0.028482653200626373
/usr/local/lib/python3.7/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=0.9642058165548099, f1=0.9610678531701891,
best_f1=0.9597315436241611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002146567218005657
step: 10, loss: 0.003425834234803915
step: 20, loss: 0.001383441500365734
step: 30, loss: 0.0012957496801391244
step: 40, loss: 0.0008947915048338473
step: 50, loss: 0.00019882505876012146
step: 60, loss: 0.001174048869870603
step: 70, loss: 0.00015114857524167746
step: 80, loss: 0.04912581294775009
step: 90, loss: 0.0022285250015556812
step: 100, loss: 0.0035996034275740385
step: 110, loss: 0.00013871454575564712
step: 120, loss: 0.009605586528778076
step: 130, loss: 8.015300409169868e-05

```

```

step: 140, loss: 0.008118657395243645
step: 150, loss: 7.193652709247544e-05
step: 160, loss: 0.00047356143477372825
step: 170, loss: 0.00010863319766940549
step: 180, loss: 0.0002763926750048995
step: 190, loss: 0.0023411852307617664
step: 200, loss: 0.0012457470875233412
step: 210, loss: 0.00023961592523846775
step: 220, loss: 0.03029169887304306
step: 230, loss: 0.001441177329979837
/usr/local/lib/python3.7/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=0.9650507328072153, f1=0.9617117117117117,
best_f1=0.9597315436241611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010842887422768399
step: 10, loss: 7.445158553309739e-05
step: 20, loss: 0.00013397315342444927
step: 30, loss: 0.00014860957162454724
step: 40, loss: 4.330814044806175e-05
step: 50, loss: 0.0008030461031012237
step: 60, loss: 0.029151037335395813
step: 70, loss: 0.00010999303049175069
step: 80, loss: 0.0002070284535875544
step: 90, loss: 0.00012254808098077774
step: 100, loss: 0.00011289092799415812
step: 110, loss: 0.0002543833397794515
step: 120, loss: 0.017435990273952484
step: 130, loss: 0.00014331471174955368
step: 140, loss: 0.017989518120884895
step: 150, loss: 0.0013707836624234915
step: 160, loss: 6.489998486358672e-05
step: 170, loss: 0.003218592843040824
step: 180, loss: 0.001237118267454207
step: 190, loss: 0.0002028374874498695
step: 200, loss: 0.00033131882082670927
step: 210, loss: 0.00010086241672979668
step: 220, loss: 0.0002359599166084081
step: 230, loss: 0.004741692915558815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.967452300785634, f1=0.9675977653631285,
best_f1=0.9675977653631285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.601577767170966e-05
step: 10, loss: 7.61892952141352e-05
step: 20, loss: 6.51151713100262e-05
step: 30, loss: 0.005727372597903013
step: 40, loss: 0.010788340121507645
step: 50, loss: 0.00023839356435928494
step: 60, loss: 7.872351125115529e-05
step: 70, loss: 0.0002178120776079595
step: 80, loss: 0.00010497759649297222
step: 90, loss: 0.00030605783103965223
step: 100, loss: 0.00011337209434714168
step: 110, loss: 7.379688031505793e-05
step: 120, loss: 8.947798778535798e-05
step: 130, loss: 5.709975812351331e-05
step: 140, loss: 5.422422691481188e-05
step: 150, loss: 5.090617560199462e-05
step: 160, loss: 5.092481660540216e-05
step: 170, loss: 0.0015941004967316985
step: 180, loss: 0.00011909494060091674
step: 190, loss: 0.00091556366533041
step: 200, loss: 7.609375461470336e-05
step: 210, loss: 5.055309520685114e-05
step: 220, loss: 6.229542486835271e-05
step: 230, loss: 0.0382038913667202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.967525195968645, f1=0.9653631284916202,
best_f1=0.9653631284916202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004298058629501611
step: 10, loss: 9.865004540188238e-05
step: 20, loss: 4.989312947145663e-05
step: 30, loss: 6.381310959113762e-05
step: 40, loss: 4.168772647972219e-05
step: 50, loss: 8.10984565760009e-05

```

```

step: 60, loss: 0.027789074927568436
step: 70, loss: 6.520210445160046e-05
step: 80, loss: 5.726470408262685e-05
step: 90, loss: 0.00021458030096255243
step: 100, loss: 3.979578468715772e-05
step: 110, loss: 0.0020495003554970026
step: 120, loss: 7.641344564035535e-05
step: 130, loss: 8.01915957708843e-05
step: 140, loss: 4.660775084630586e-05
step: 150, loss: 3.391736390767619e-05
step: 160, loss: 0.025702854618430138
step: 170, loss: 8.144485764205456e-05
step: 180, loss: 4.1967115976149216e-05
step: 190, loss: 4.413393253344111e-05
step: 200, loss: 0.010923651978373528
step: 210, loss: 5.861551471753046e-05
step: 220, loss: 0.01030303630977869
step: 230, loss: 7.475598249584436e-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=0.963882618510158, f1=0.9659090909090909,
best_f1=0.9653631284916202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000261381093878299
step: 10, loss: 6.536878936458379e-05
step: 20, loss: 7.512736920034513e-05
step: 30, loss: 0.014107768423855305
step: 40, loss: 0.0005998507258482277
step: 50, loss: 9.062428580364212e-05
step: 60, loss: 4.948249625158496e-05
step: 70, loss: 9.815376688493416e-05
step: 80, loss: 7.033885776763782e-05
step: 90, loss: 3.125422517769039e-05
step: 100, loss: 6.595281593035907e-05
step: 110, loss: 5.5156800954137e-05
step: 120, loss: 0.0019279003608971834
step: 130, loss: 0.00011314354924252257
step: 140, loss: 0.001792051480151713
step: 150, loss: 0.00046816118992865086
step: 160, loss: 6.555338768521324e-05
step: 170, loss: 5.262747436063364e-05
step: 180, loss: 0.0002510450140107423
step: 190, loss: 0.0023894445039331913
step: 200, loss: 0.00010925607057288289
step: 210, loss: 0.017680229619145393

```

```

step: 220, loss: 7.555228512501344e-05
step: 230, loss: 4.1117229557130486e-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=0.9650507328072153, f1=0.9626274065685164,
best_f1=0.9653631284916202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.175189577741548e-05
step: 10, loss: 3.1786046747583896e-05
step: 20, loss: 0.005138102453202009
step: 30, loss: 0.00014532089699059725
step: 40, loss: 3.3883559808600694e-05
step: 50, loss: 0.0003931844257749617
step: 60, loss: 0.00010884247603826225
step: 70, loss: 3.945280332118273e-05
step: 80, loss: 7.389693928416818e-05
step: 90, loss: 0.00011089040344813839
step: 100, loss: 0.004911540541797876
step: 110, loss: 8.799002534942701e-05
step: 120, loss: 5.303479701979086e-05
step: 130, loss: 7.940318755572662e-05
step: 140, loss: 4.418200842337683e-05
step: 150, loss: 4.880244887317531e-05
step: 160, loss: 4.141417593928054e-05
step: 170, loss: 0.0006164521910250187
step: 180, loss: 5.8534726122161373e-05
step: 190, loss: 0.00033344700932502747
step: 200, loss: 7.23158082109876e-05
step: 210, loss: 0.00023301872715819627
step: 220, loss: 6.252994353417307e-05
step: 230, loss: 2.165847399737686e-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=0.963963963963964, f1=0.9617977528089887,
best_f1=0.9653631284916202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4207943826913834e-05
step: 10, loss: 5.1089024054817855e-05
step: 20, loss: 0.026283979415893555
step: 30, loss: 7.148674194468185e-05
step: 40, loss: 0.00011053874186472967
step: 50, loss: 5.420706293080002e-05

```

```

step: 60, loss: 5.714026337955147e-05
step: 70, loss: 4.505868855630979e-05
step: 80, loss: 5.0824484787881374e-05
step: 90, loss: 2.740619493124541e-05
step: 100, loss: 8.328547119162977e-05
step: 110, loss: 3.986258161603473e-05
step: 120, loss: 9.006318578030914e-05
step: 130, loss: 0.00010289095371263102
step: 140, loss: 0.0001083505485439673
step: 150, loss: 8.583159069530666e-05
step: 160, loss: 9.280692756874487e-05
step: 170, loss: 3.281836325186305e-05
step: 180, loss: 6.679004582110792e-05
step: 190, loss: 0.00011402558448025957
step: 200, loss: 3.217041012248956e-05
step: 210, loss: 0.00022758357226848602
step: 220, loss: 8.719313336769119e-05
step: 230, loss: 5.445328133646399e-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=0.9649717514124293, f1=0.963882618510158,
best_f1=0.9653631284916202
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 [00:08, 286.95it/s]
load_f1 = 0.967525195968645
real_f1 = 0.9686800894854586
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 [00:12, 351.62it/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 distilbert \
  --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 distilbert \
  --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: 0.8000226616859436
step: 10, loss: 0.40837061405181885
step: 20, loss: 0.4976736307144165
step: 30, loss: 0.44980233907699585
step: 40, loss: 0.4379374384880066
step: 50, loss: 0.224131241440773
```

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

```
step: 60, loss: 0.3235783278942108
step: 70, loss: 0.06913358718156815
step: 80, loss: 0.11769244074821472
step: 90, loss: 0.19253437221050262
step: 100, loss: 0.2856212258338928
```

```
step: 110, loss: 0.09063132852315903
step: 120, loss: 0.14985787868499756
step: 130, loss: 0.06828735023736954
step: 140, loss: 0.1691695898771286
step: 150, loss: 0.11401953548192978
step: 160, loss: 0.17375481128692627
step: 170, loss: 0.26627978682518005
step: 180, loss: 0.1409723311662674
step: 190, loss: 0.07425282895565033
step: 200, loss: 0.12708987295627594
step: 210, loss: 0.1337747722864151
step: 220, loss: 0.20476676523685455
step: 230, loss: 0.15341679751873016
step: 240, loss: 0.08568456023931503
step: 250, loss: 0.07131136208772659
step: 260, loss: 0.018310001119971275
step: 270, loss: 0.019451381638646126
step: 280, loss: 0.2759746015071869
step: 290, loss: 0.02027282491326332
step: 300, loss: 0.18212686479091644
step: 310, loss: 0.11783904582262039
step: 320, loss: 0.09764725714921951
step: 330, loss: 0.16329623758792877
step: 340, loss: 0.24021261930465698
step: 350, loss: 0.08812611550092697
step: 360, loss: 0.06569600850343704
step: 370, loss: 0.08837087452411652
step: 380, loss: 0.2883330285549164
step: 390, loss: 0.022044850513339043
step: 400, loss: 0.07304628938436508
step: 410, loss: 0.13452085852622986
step: 420, loss: 0.0378718376159668
step: 430, loss: 0.045574091374874115
step: 440, loss: 0.2071564942598343
step: 450, loss: 0.06274071335792542
step: 460, loss: 0.050669025629758835
step: 470, loss: 0.3241581916809082
step: 480, loss: 0.1577594131231308
step: 490, loss: 0.1291217803955078
step: 500, loss: 0.12943655252456665
step: 510, loss: 0.1031670942902565
step: 520, loss: 0.13197669386863708
step: 530, loss: 0.12438464164733887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9018492176386912, f1=0.9080622347949081,
best_f1=0.9080622347949081
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07643730938434601
step: 10, loss: 0.1400427371263504
step: 20, loss: 0.14375831186771393
step: 30, loss: 0.02105908840894699
step: 40, loss: 0.009352910332381725
step: 50, loss: 0.06960909813642502
step: 60, loss: 0.21181116998195648
step: 70, loss: 0.10269960761070251
step: 80, loss: 0.026648113504052162
step: 90, loss: 0.04537755250930786
step: 100, loss: 0.21230505406856537
step: 110, loss: 0.09620913863182068
step: 120, loss: 0.060802657157182693
step: 130, loss: 0.08264323323965073
step: 140, loss: 0.13379763066768646
step: 150, loss: 0.07393372058868408
step: 160, loss: 0.0532270111143589
step: 170, loss: 0.12431208789348602
step: 180, loss: 0.025970207527279854
step: 190, loss: 0.07545117288827896
step: 200, loss: 0.02358979359269142
step: 210, loss: 0.022138018161058426
step: 220, loss: 0.14686952531337738
step: 230, loss: 0.08354270458221436
step: 240, loss: 0.15312889218330383
step: 250, loss: 0.11393848806619644
step: 260, loss: 0.022329263389110565
step: 270, loss: 0.0804230123758316
step: 280, loss: 0.3626182973384857
step: 290, loss: 0.0872211903333664
step: 300, loss: 0.042380284518003464
step: 310, loss: 0.05052199214696884
step: 320, loss: 0.15065120160579681
step: 330, loss: 0.03347950428724289
step: 340, loss: 0.018515750765800476
step: 350, loss: 0.04610852897167206
step: 360, loss: 0.130979984998703
step: 370, loss: 0.021950246766209602
step: 380, loss: 0.10372214019298553
step: 390, loss: 0.042155541479587555
step: 400, loss: 0.1343553066253662

```

```

step: 410, loss: 0.027635766193270683
step: 420, loss: 0.10523892939090729
step: 430, loss: 0.026841875165700912
step: 440, loss: 0.01988452859222889
step: 450, loss: 0.015347050502896309
step: 460, loss: 0.3852153420448303
step: 470, loss: 0.09518644213676453
step: 480, loss: 0.19301581382751465
step: 490, loss: 0.06358662247657776
step: 500, loss: 0.03718562424182892
step: 510, loss: 0.08987286686897278
step: 520, loss: 0.01568070612847805
step: 530, loss: 0.12021682411432266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9209171736078615, f1=0.9155844155844156,
best_f1=0.9155844155844156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0724889263510704
step: 10, loss: 0.10304751247167587
step: 20, loss: 0.273441880941391
step: 30, loss: 0.10439848154783249
step: 40, loss: 0.03208480775356293
step: 50, loss: 0.06452810019254684
step: 60, loss: 0.04663675278425217
step: 70, loss: 0.04697834700345993
step: 80, loss: 0.03023933432996273
step: 90, loss: 0.24125991761684418
step: 100, loss: 0.07968377321958542
step: 110, loss: 0.031678780913352966
step: 120, loss: 0.07311864197254181
step: 130, loss: 0.04345620796084404
step: 140, loss: 0.03477337211370468
step: 150, loss: 0.04138966277241707
step: 160, loss: 0.012193309143185616
step: 170, loss: 0.005132784601300955
step: 180, loss: 0.01606191322207451
step: 190, loss: 0.008020911365747452
step: 200, loss: 0.0769331157207489
step: 210, loss: 0.10223022103309631
step: 220, loss: 0.008753609843552113

```

```

step: 230, loss: 0.10941697657108307
step: 240, loss: 0.03232021629810333
step: 250, loss: 0.02456609532237053
step: 260, loss: 0.006407481152564287
step: 270, loss: 0.013658411800861359
step: 280, loss: 0.055416710674762726
step: 290, loss: 0.06855684518814087
step: 300, loss: 0.06104755774140358
step: 310, loss: 0.14229564368724823
step: 320, loss: 0.23963984847068787
step: 330, loss: 0.009932623244822025
step: 340, loss: 0.010711109265685081
step: 350, loss: 0.0943567231297493
step: 360, loss: 0.02409006468951702
step: 370, loss: 0.014033551327884197
step: 380, loss: 0.049785811454057693
step: 390, loss: 0.02921462245285511
step: 400, loss: 0.10332409292459488
step: 410, loss: 0.012919233180582523
step: 420, loss: 0.038577694445848465
step: 430, loss: 0.01839599758386612
step: 440, loss: 0.043049126863479614
step: 450, loss: 0.10673854500055313
step: 460, loss: 0.10573925822973251
step: 470, loss: 0.019610228016972542
step: 480, loss: 0.020954487845301628
step: 490, loss: 0.13209408521652222
step: 500, loss: 0.10311675816774368
step: 510, loss: 0.007325653452426195
step: 520, loss: 0.006913961376994848
step: 530, loss: 0.01649552769958973
/usr/local/lib/python3.7/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=0.917516218721038, f1=0.9157407407407409,
best_f1=0.9155844155844156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02323302812874317
step: 10, loss: 0.05224825441837311
step: 20, loss: 0.037730276584625244
step: 30, loss: 0.024168923497200012
step: 40, loss: 0.014884968288242817
step: 50, loss: 0.01232975535094738
step: 60, loss: 0.00319151533767581
step: 70, loss: 0.0021530508529394865
step: 80, loss: 0.030348174273967743

```

```
step: 90, loss: 0.03688390180468559
step: 100, loss: 0.008727189153432846
step: 110, loss: 0.052011072635650635
step: 120, loss: 0.012896421365439892
step: 130, loss: 0.007525905966758728
step: 140, loss: 0.0635102391242981
step: 150, loss: 0.0010712961666285992
step: 160, loss: 0.010619021952152252
step: 170, loss: 0.04466945305466652
step: 180, loss: 0.07083941996097565
step: 190, loss: 0.0550549179315567
step: 200, loss: 0.03983564302325249
step: 210, loss: 0.02026781439781189
step: 220, loss: 0.006701957434415817
step: 230, loss: 0.19864042103290558
step: 240, loss: 0.017603011801838875
step: 250, loss: 0.008458174765110016
step: 260, loss: 0.1568283885717392
step: 270, loss: 0.044781189411878586
step: 280, loss: 0.013711811043322086
step: 290, loss: 0.07315816730260849
step: 300, loss: 0.002562585985288024
step: 310, loss: 0.04313419759273529
step: 320, loss: 0.018897587433457375
step: 330, loss: 0.15895166993141174
step: 340, loss: 0.19438551366329193
step: 350, loss: 0.08769700676202774
step: 360, loss: 0.029441600665450096
step: 370, loss: 0.09921081364154816
step: 380, loss: 0.001466191723011434
step: 390, loss: 0.031092753633856773
step: 400, loss: 0.051744718104600906
step: 410, loss: 0.01610541343688965
step: 420, loss: 0.037261445075273514
step: 430, loss: 0.026600036770105362
step: 440, loss: 0.0544196292757988
step: 450, loss: 0.05433996394276619
step: 460, loss: 0.002519944915547967
step: 470, loss: 0.04285745322704315
step: 480, loss: 0.0018791293259710073
step: 490, loss: 0.011555474251508713
step: 500, loss: 0.02503420040011406
step: 510, loss: 0.021944282576441765
step: 520, loss: 0.07803978770971298
step: 530, loss: 0.0038805522490292788
/usr/local/lib/python3.7/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=0.9152073732718895, f1=0.912665752171925,
best_f1=0.9155844155844156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006466939579695463
step: 10, loss: 0.028362810611724854
step: 20, loss: 0.03250846266746521
step: 30, loss: 0.00892863143235445
step: 40, loss: 0.04027177020907402
step: 50, loss: 0.007288915570825338
step: 60, loss: 0.05464752018451691
step: 70, loss: 0.00046178465709090233
step: 80, loss: 0.004146668594330549
step: 90, loss: 0.0012268080608919263
step: 100, loss: 0.0016047528479248285
step: 110, loss: 0.060358524322509766
step: 120, loss: 0.007377467583864927
step: 130, loss: 0.0015445119934156537
step: 140, loss: 0.003715969854965806
step: 150, loss: 0.016665711998939514
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.049569349735975266
step: 170, loss: 0.04538077861070633
step: 180, loss: 0.008476927876472473
step: 190, loss: 0.0037866856437176466
step: 200, loss: 0.017300814390182495
step: 210, loss: 0.0026029464788734913
step: 220, loss: 0.004058867692947388
step: 230, loss: 0.003111253259703517
step: 240, loss: 0.0022550830617547035
step: 250, loss: 0.00040941929910331964
step: 260, loss: 0.024837447330355644
step: 270, loss: 0.05745266005396843
step: 280, loss: 0.03452889248728752
step: 290, loss: 0.0014043367700651288
step: 300, loss: 0.014419957995414734
step: 310, loss: 0.06939512491226196
step: 320, loss: 0.021551862359046936
step: 330, loss: 0.004434189759194851
step: 340, loss: 0.016552183777093887
step: 350, loss: 0.000865231326315552
step: 360, loss: 0.0239464920014143
step: 370, loss: 0.012970766052603722
step: 380, loss: 0.013598363846540451
step: 390, loss: 0.0015105954371392727
step: 400, loss: 0.010098912753164768
step: 410, loss: 0.010979482904076576

```

```

step: 420, loss: 0.0016738856211304665
step: 430, loss: 0.024331580847501755
step: 440, loss: 0.014098242856562138
step: 450, loss: 0.09490609169006348
step: 460, loss: 0.1117643192410469
step: 470, loss: 0.004283765330910683
step: 480, loss: 0.019564494490623474
step: 490, loss: 0.08582032471895218
step: 500, loss: 0.04198735952377319
step: 510, loss: 0.009505750611424446
step: 520, loss: 0.00683105643838644
step: 530, loss: 0.016866909340023994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9246607393542348, f1=0.9154929577464789,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002880742773413658
step: 10, loss: 0.019500281661748886
step: 20, loss: 0.00047601156984455884
step: 30, loss: 0.0054455650970339775
step: 40, loss: 0.02131788246333599
step: 50, loss: 0.0015470562502741814
step: 60, loss: 0.02679385244846344
step: 70, loss: 0.010790450498461723
step: 80, loss: 0.0009721017559058964
step: 90, loss: 0.03625477850437164
step: 100, loss: 0.00576743995770812
step: 110, loss: 0.01027825940400362
step: 120, loss: 0.013288321904838085
step: 130, loss: 0.0017982349963858724
step: 140, loss: 0.0009081689058803022
step: 150, loss: 0.003717460436746478
step: 160, loss: 0.0017821459332481027
step: 170, loss: 0.0005039884126745164
step: 180, loss: 0.012561422772705555
step: 190, loss: 0.28532373905181885
step: 200, loss: 0.0009987769881263375
step: 210, loss: 0.062121886759996414
step: 220, loss: 0.0028503770008683205
step: 230, loss: 0.002994804410263896

```



```

step: 240, loss: 0.0025654856581240892
step: 250, loss: 0.005429232958704233
step: 260, loss: 0.07780281454324722
step: 270, loss: 0.0022556076291948557
step: 280, loss: 0.01817280799150467
step: 290, loss: 0.002511025872081518
step: 300, loss: 0.0021672395523637533
step: 310, loss: 0.0033502248115837574
step: 320, loss: 0.04382263869047165
step: 330, loss: 0.006284893956035376
step: 340, loss: 0.028319628909230232
step: 350, loss: 0.16170169413089752
step: 360, loss: 0.0006380091072060168
step: 370, loss: 0.010509451851248741
step: 380, loss: 0.003127865493297577
step: 390, loss: 0.16774307191371918
step: 400, loss: 0.0006099808379076421
step: 410, loss: 0.00040330164483748376
step: 420, loss: 0.08995050191879272
step: 430, loss: 0.15072020888328552
step: 440, loss: 0.06769192218780518
step: 450, loss: 0.0038767443038523197
step: 460, loss: 0.0032925563864409924
step: 470, loss: 0.0817830041050911
step: 480, loss: 0.0020750670228153467
step: 490, loss: 0.006985748652368784
step: 500, loss: 0.0011126062599942088
step: 510, loss: 0.0014525556471198797
step: 520, loss: 0.0009625544189475477
step: 530, loss: 0.0007815533899702132
/usr/local/lib/python3.7/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=0.9118340405469118, f1=0.9150812064965197,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014219058211892843
step: 10, loss: 0.030045976862311363
step: 20, loss: 0.0038658075500279665
step: 30, loss: 0.009469615295529366
step: 40, loss: 0.019567709416151047
step: 50, loss: 0.003697648411616683
step: 60, loss: 0.014026318676769733
step: 70, loss: 0.036344315856695175
step: 80, loss: 0.003541344776749611
step: 90, loss: 0.0031494449358433485

```

```
step: 100, loss: 0.004957453813403845
step: 110, loss: 0.0275123193860054
step: 120, loss: 0.00046597045729868114
step: 130, loss: 0.00034855713602155447
step: 140, loss: 0.01635022833943367
step: 150, loss: 0.003601571312174201
step: 160, loss: 0.0037793226074427366
step: 170, loss: 0.09296078234910965
step: 180, loss: 0.0013248661998659372
step: 190, loss: 0.00024050667707342654
step: 200, loss: 0.004276927560567856
step: 210, loss: 0.0012536441208794713
step: 220, loss: 0.000525006209500134
step: 230, loss: 0.00033125514164566994
step: 240, loss: 0.002568742958828807
step: 250, loss: 0.0007106020348146558
step: 260, loss: 0.007169822230935097
step: 270, loss: 0.0010150879388675094
step: 280, loss: 0.01255623996257782
step: 290, loss: 0.004395169205963612
step: 300, loss: 0.007881319150328636
step: 310, loss: 0.022887421771883965
step: 320, loss: 0.05348159000277519
step: 330, loss: 0.0002641906321514398
step: 340, loss: 0.011637002229690552
step: 350, loss: 0.004771876148879528
step: 360, loss: 0.0011699867900460958
step: 370, loss: 0.00042546793702058494
step: 380, loss: 0.00023654663527850062
step: 390, loss: 0.000168391372426413
step: 400, loss: 0.00015609986439812928
step: 410, loss: 0.00024064219905994833
step: 420, loss: 0.0004796126449946314
step: 430, loss: 0.027886351570487022
step: 440, loss: 0.0011623702012002468
step: 450, loss: 0.00022189687297213823
step: 460, loss: 0.011251276358962059
step: 470, loss: 0.011489210650324821
step: 480, loss: 0.003694843500852585
step: 490, loss: 0.00033122923923656344
step: 500, loss: 0.0004958570934832096
step: 510, loss: 0.008333366364240646
step: 520, loss: 0.004312792792916298
step: 530, loss: 0.0005597705021500587
/usr/local/lib/python3.7/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=0.9120521172638437, f1=0.9163972286374134,
```

```
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021675440948456526
step: 10, loss: 0.0005808043060824275
step: 20, loss: 0.006987337488681078
step: 30, loss: 0.013824861496686935
step: 40, loss: 0.007501524873077869
step: 50, loss: 0.00046613457379862666
step: 60, loss: 0.00025614103651605546
step: 70, loss: 0.008395787328481674
step: 80, loss: 0.0018208121182397008
step: 90, loss: 0.002588003408163786
step: 100, loss: 0.001991485944017768
step: 110, loss: 0.007344452664256096
step: 120, loss: 0.0030872051138430834
step: 130, loss: 0.036088258028030396
step: 140, loss: 0.0004889328847639263
step: 150, loss: 0.010694021359086037
step: 160, loss: 0.010547497309744358
step: 170, loss: 0.0005169657524675131
step: 180, loss: 0.0012809372274205089
step: 190, loss: 0.0010976551566272974
step: 200, loss: 0.0015016965335235
step: 210, loss: 0.0009128224919550121
step: 220, loss: 0.003412848338484764
step: 230, loss: 0.0024557916913181543
step: 240, loss: 0.03589111939072609
step: 250, loss: 0.003062186064198613
step: 260, loss: 0.10027355700731277
step: 270, loss: 0.0010195579379796982
step: 280, loss: 0.00023313051497098058
step: 290, loss: 0.12293966114521027
step: 300, loss: 0.0015892356168478727
step: 310, loss: 0.02848937176167965
step: 320, loss: 0.0027570046950131655
step: 330, loss: 0.04028626158833504
step: 340, loss: 0.0037365825846791267
step: 350, loss: 0.006161041092127562
step: 360, loss: 0.00037323907599784434
step: 370, loss: 0.006109252572059631
step: 380, loss: 0.030624108389019966
step: 390, loss: 0.0036437432281672955
step: 400, loss: 0.032043542712926865
step: 410, loss: 0.0025127683766186237
step: 420, loss: 0.02874147891998291
step: 430, loss: 0.0012236350448802114
```

```

step: 440, loss: 0.005711089354008436
step: 450, loss: 0.002605077112093568
step: 460, loss: 0.0008740506018511951
step: 470, loss: 0.06775031238794327
step: 480, loss: 0.002153735840693116
step: 490, loss: 0.0005273165297694504
step: 500, loss: 0.01393511239439249
step: 510, loss: 0.0019348637433722615
step: 520, loss: 0.003054215107113123
step: 530, loss: 0.0037972733844071627
/usr/local/lib/python3.7/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=0.9133858267716536, f1=0.9165514061779623,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006124000530689955
step: 10, loss: 0.0006173504516482353
step: 20, loss: 0.0006210512947291136
step: 30, loss: 0.00017397169722244143
step: 40, loss: 0.018797753378748894
step: 50, loss: 0.001943318173289299
step: 60, loss: 0.0018740170635282993
step: 70, loss: 0.004533517174422741
step: 80, loss: 0.00019466331286821514
step: 90, loss: 0.0005525480955839157
step: 100, loss: 0.005960378330200911
step: 110, loss: 0.00486814696341753
step: 120, loss: 0.0122081208974123
step: 130, loss: 0.0016829310916364193
step: 140, loss: 0.0011436698259785771
step: 150, loss: 0.002814274514093995
step: 160, loss: 0.0004455498419702053
step: 170, loss: 0.00029928924050182104
step: 180, loss: 0.000440722971688956
step: 190, loss: 0.002491411752998829
step: 200, loss: 0.0004968594294041395
step: 210, loss: 0.000192440245882608
step: 220, loss: 0.010788426734507084
step: 230, loss: 0.0015057845739647746
step: 240, loss: 0.00020235929696355015
step: 250, loss: 0.02829120121896267
step: 260, loss: 0.004488281439989805
step: 270, loss: 0.01010490208864212
step: 280, loss: 0.00020783163199666888
step: 290, loss: 4.895096208201721e-05

```

```

step: 300, loss: 0.002023997250944376
step: 310, loss: 0.00011096634261775762
step: 320, loss: 4.936481127515435e-05
step: 330, loss: 0.0003658513305708766
step: 340, loss: 0.000479514361359179
step: 350, loss: 0.00016037735622376204
step: 360, loss: 0.027799535542726517
step: 370, loss: 0.020511507987976074
step: 380, loss: 0.00014018986257724464
step: 390, loss: 0.00017504738934803754
step: 400, loss: 0.0018380164401605725
step: 410, loss: 0.031741414219141006
step: 420, loss: 0.0021023056469857693
step: 430, loss: 4.885785892838612e-05
step: 440, loss: 0.00025048357201740146
step: 450, loss: 0.0002937505778390914
step: 460, loss: 0.0015257364138960838
step: 470, loss: 9.371792111778632e-05
step: 480, loss: 0.00013266646419651806
step: 490, loss: 0.002562924986705184
step: 500, loss: 0.031356748193502426
step: 510, loss: 0.024584993720054626
step: 520, loss: 0.003489942289888859
step: 530, loss: 0.00048275015433318913
/usr/local/lib/python3.7/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=0.9178966789667896, f1=0.9132629646626893,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00954581517726183
step: 10, loss: 0.000301752908853814
step: 20, loss: 0.005392353981733322
step: 30, loss: 0.001052378909662366
step: 40, loss: 0.0003649891878012568
step: 50, loss: 0.0023167275357991457
step: 60, loss: 0.0003469572402536869
step: 70, loss: 0.0009342951234430075
step: 80, loss: 0.0005005949060432613
step: 90, loss: 0.0009295566705986857
step: 100, loss: 0.0002744764497037977
step: 110, loss: 0.0004486734396778047
step: 120, loss: 0.0003934649284929037
step: 130, loss: 0.0035179436672478914
step: 140, loss: 0.00047206462477333844
step: 150, loss: 0.03525654599070549

```

```

step: 160, loss: 0.000924475200008601
step: 170, loss: 0.0004485218378249556
step: 180, loss: 0.003684171475470066
step: 190, loss: 0.0026232271920889616
step: 200, loss: 0.020798513665795326
step: 210, loss: 0.00042421382386237383
step: 220, loss: 0.0030042820144444704
step: 230, loss: 0.000869997194968164
step: 240, loss: 0.0011474085040390491
step: 250, loss: 0.0006416208925656974
step: 260, loss: 0.0033382298424839973
step: 270, loss: 0.005254996474832296
step: 280, loss: 7.49087193980813e-05
step: 290, loss: 0.0037763849832117558
step: 300, loss: 0.0027847932651638985
step: 310, loss: 0.001194479176774621
step: 320, loss: 0.0005572015070356429
step: 330, loss: 0.0030313178431242704
step: 340, loss: 8.117692777886987e-05
step: 350, loss: 8.961361163528636e-05
step: 360, loss: 0.014121823012828827
step: 370, loss: 0.06866846233606339
step: 380, loss: 0.012384343892335892
step: 390, loss: 5.4311232815962285e-05
step: 400, loss: 0.023651542142033577
step: 410, loss: 0.001444765948690474
step: 420, loss: 0.008555521257221699
step: 430, loss: 0.0002637863508425653
step: 440, loss: 3.0978178983787075e-05
step: 450, loss: 0.0016127409180626273
step: 460, loss: 0.0008763546356931329
step: 470, loss: 0.00018307447317056358
step: 480, loss: 0.005954371765255928
step: 490, loss: 0.11345706135034561
step: 500, loss: 0.013857241719961166
step: 510, loss: 0.006586684845387936
step: 520, loss: 0.009096825495362282
step: 530, loss: 0.0002907453163061291
/usr/local/lib/python3.7/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=0.9140733859730609, f1=0.9136029411764707,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001146246213465929
step: 10, loss: 0.0026522218249738216

```

step: 20, loss: 9.349596803076565e-05
step: 30, loss: 0.011009838432073593
step: 40, loss: 0.00016137692728079855
step: 50, loss: 0.000638654048088938
step: 60, loss: 0.00035711631062440574
step: 70, loss: 2.330497773073148e-05
step: 80, loss: 0.00019846776558551937
step: 90, loss: 0.0004450620908755809
step: 100, loss: 9.086859063245356e-05
step: 110, loss: 2.578972271294333e-05
step: 120, loss: 5.715589577448554e-05
step: 130, loss: 6.71920643071644e-05
step: 140, loss: 0.0014978667022660375
step: 150, loss: 0.003228049958124757
step: 160, loss: 0.013687556609511375
step: 170, loss: 0.004703393671661615
step: 180, loss: 0.00011590234498726204
step: 190, loss: 0.0005389800644479692
step: 200, loss: 0.0001584626588737592
step: 210, loss: 0.0008155949180945754
step: 220, loss: 0.0009559373720549047
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 230, loss: 0.004358926322311163
step: 240, loss: 2.7786114515038207e-05
step: 250, loss: 0.00013090237916912884
step: 260, loss: 2.024655623245053e-05
step: 270, loss: 0.0008088051108643413
step: 280, loss: 0.0021245896350592375
step: 290, loss: 0.0015014547389000654
step: 300, loss: 0.002124116290360689
step: 310, loss: 0.017695175483822823
step: 320, loss: 0.018070176243782043
step: 330, loss: 0.0013250974006950855
step: 340, loss: 0.00014857303176540881
step: 350, loss: 0.0005331971915438771
step: 360, loss: 0.0038788262754678726
step: 370, loss: 0.00011422715033404529
step: 380, loss: 0.0027436562813818455
step: 390, loss: 0.12033375352621078
step: 400, loss: 4.701621946878731e-05
step: 410, loss: 0.0008026850991882384
step: 420, loss: 0.0010553468018770218
step: 430, loss: 0.005389948841184378
step: 440, loss: 0.0010540034854784608
step: 450, loss: 0.0012326923897489905
step: 460, loss: 0.001382904825732112
step: 470, loss: 0.0005507101304829121
step: 480, loss: 0.00018527124484535307

```
step: 490, loss: 0.0002702103811316192
step: 500, loss: 0.0009473804384469986
step: 510, loss: 0.000326831650454551
step: 520, loss: 4.9434471293352544e-05
step: 530, loss: 0.0013430211693048477
/usr/local/lib/python3.7/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=0.9146110056925996, f1=0.912330051570558,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006367792375385761
step: 10, loss: 0.018591683357954025
step: 20, loss: 4.336492202128284e-05
step: 30, loss: 0.003592114895582199
step: 40, loss: 0.00042921077692881227
step: 50, loss: 0.000276372826192528
step: 60, loss: 0.13699842989444733
step: 70, loss: 0.000348039495293051
step: 80, loss: 0.0008280235342681408
step: 90, loss: 0.006600192282348871
step: 100, loss: 0.003929764963686466
step: 110, loss: 0.004214849788695574
step: 120, loss: 0.004074994008988142
step: 130, loss: 0.0006509466329589486
step: 140, loss: 0.007109232246875763
step: 150, loss: 0.004521610680967569
step: 160, loss: 0.04803762212395668
step: 170, loss: 0.00889821071177721
step: 180, loss: 0.00032619148259982467
step: 190, loss: 0.008884434588253498
step: 200, loss: 0.0004558721266221255
step: 210, loss: 0.00038541154935956
step: 220, loss: 0.00021368314628489316
step: 230, loss: 0.009440300054848194
step: 240, loss: 0.0004727437917608768
step: 250, loss: 0.0029017433989793062
step: 260, loss: 0.00036291012656874955
step: 270, loss: 0.0001981997484108433
step: 280, loss: 5.504941873368807e-05
step: 290, loss: 0.0017005972331389785
step: 300, loss: 0.0004182069096714258
step: 310, loss: 0.0009079402079805732
step: 320, loss: 0.000355343654518947
step: 330, loss: 0.00015702092787250876
step: 340, loss: 0.00042162626050412655
```



```
step: 350, loss: 0.029133141040802002
step: 360, loss: 0.0008771239081397653
step: 370, loss: 0.0016396265709772706
step: 380, loss: 0.0009765059221535921
step: 390, loss: 0.00010149455920327455
step: 400, loss: 0.002024887129664421
step: 410, loss: 0.007386075332760811
step: 420, loss: 0.0014368194388225675
step: 430, loss: 0.0014306543162092566
step: 440, loss: 0.000567909621167928
step: 450, loss: 0.01782168634235859
step: 460, loss: 0.022741733118891716
step: 470, loss: 0.003152469638735056
step: 480, loss: 0.00017513845523353666
step: 490, loss: 0.0014743634965270758
step: 500, loss: 0.00492476113140583
step: 510, loss: 0.00014424038818106055
step: 520, loss: 0.028283286839723587
step: 530, loss: 0.0002954781521111727
/usr/local/lib/python3.7/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=0.922071861875875, f1=0.9174397031539888,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003596313064917922
step: 10, loss: 0.00019660881662275642
step: 20, loss: 0.02361847460269928
step: 30, loss: 0.031353551894426346
step: 40, loss: 0.00016265704471152276
step: 50, loss: 0.0002130211505573243
step: 60, loss: 5.599388896371238e-05
step: 70, loss: 0.0008514462970197201
step: 80, loss: 0.004168620798736811
step: 90, loss: 0.0037978750187903643
step: 100, loss: 6.486952770501375e-05
step: 110, loss: 0.005105890799313784
step: 120, loss: 0.0010094335302710533
step: 130, loss: 0.004165142774581909
step: 140, loss: 0.005316440016031265
step: 150, loss: 9.948501246981323e-05
step: 160, loss: 0.00023110947222448885
step: 170, loss: 0.0010175463976338506
step: 180, loss: 0.0006666244589723647
step: 190, loss: 0.00047582387924194336
step: 200, loss: 0.002555603627115488
```

```

step: 210, loss: 0.0026037723291665316
step: 220, loss: 0.0017049937741830945
step: 230, loss: 0.001144773093983531
step: 240, loss: 3.8718477298971266e-05
step: 250, loss: 0.006218325812369585
step: 260, loss: 0.0044679222628474236
step: 270, loss: 0.0009837434627115726
step: 280, loss: 0.005628902465105057
step: 290, loss: 0.0010552455205470324
step: 300, loss: 0.0005219351151026785
step: 310, loss: 0.0017894419142976403
step: 320, loss: 8.158569107763469e-05
step: 330, loss: 0.0008108928450383246
step: 340, loss: 0.000198462963453494
step: 350, loss: 0.06375893205404282
step: 360, loss: 0.0004393197887111455
step: 370, loss: 0.028444910421967506
step: 380, loss: 0.00475294841453433
step: 390, loss: 0.00015658873599022627
step: 400, loss: 0.00011262338375672698
step: 410, loss: 0.0004167848965153098
step: 420, loss: 0.00014157289115246385
step: 430, loss: 5.6547403801232576e-05
step: 440, loss: 0.00034428227809257805
step: 450, loss: 0.001473437063395977
step: 460, loss: 1.3962222510599531e-05
step: 470, loss: 0.017675600945949554
step: 480, loss: 0.0006525891367346048
step: 490, loss: 0.0018663323717191815
step: 500, loss: 0.008546222932636738
step: 510, loss: 0.0035082437098026276
step: 520, loss: 0.00011850916780531406
step: 530, loss: 0.00035656685940921307
/usr/local/lib/python3.7/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=0.9177126917712692, f1=0.9111008751727315,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020611388026736677
step: 10, loss: 0.0024892077781260014
step: 20, loss: 0.0002528147888369858
step: 30, loss: 0.00023003101523499936
step: 40, loss: 0.005979639012366533
step: 50, loss: 0.0004427130916155875
step: 60, loss: 1.965416049642954e-05

```

step: 70, loss: 7.900034688645974e-05
step: 80, loss: 2.843737820512615e-05
step: 90, loss: 3.336430745548569e-05
step: 100, loss: 0.029411358758807182
step: 110, loss: 0.00010801263124449179
step: 120, loss: 0.00019975542090833187
step: 130, loss: 0.0002809388388413936
step: 140, loss: 0.02907605655491352
step: 150, loss: 0.00035182450665161014
step: 160, loss: 0.0002390359586570412
step: 170, loss: 0.048473842442035675
step: 180, loss: 0.000344959000358358
step: 190, loss: 0.0006759028183296323
step: 200, loss: 6.322279659798369e-05
step: 210, loss: 0.00020937068620696664
step: 220, loss: 0.0010059659834951162
step: 230, loss: 0.0005607504863291979
step: 240, loss: 0.0008783232769928873
step: 250, loss: 0.0250601414591074
step: 260, loss: 7.430026744259521e-05
step: 270, loss: 0.0015848568873479962
step: 280, loss: 0.00015118987357709557
step: 290, loss: 0.00026820111088454723
step: 300, loss: 0.0003710088785737753
step: 310, loss: 0.0017510310281068087
step: 320, loss: 0.0014873137697577477
step: 330, loss: 0.005132402293384075
step: 340, loss: 0.0003201159997843206
step: 350, loss: 0.0007266110042110085
step: 360, loss: 0.0014889773447066545
step: 370, loss: 0.11882053315639496
step: 380, loss: 0.00016316212713718414
step: 390, loss: 0.0015930355293676257
step: 400, loss: 3.432963421801105e-05
step: 410, loss: 8.719140168977901e-05
step: 420, loss: 0.0005423629190772772
step: 430, loss: 0.00010432874114485458
step: 440, loss: 0.0002691948611754924
step: 450, loss: 0.0001604632125236094
step: 460, loss: 0.0005641980678774416
step: 470, loss: 0.0002619391307234764
step: 480, loss: 8.277825691038743e-05
step: 490, loss: 0.0005397518980316818
step: 500, loss: 0.00025676455697976053
step: 510, loss: 0.011462409049272537
step: 520, loss: 0.0002966726024169475
step: 530, loss: 0.00033775740303099155
/usr/local/lib/python3.7/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=0.9149036201222379, f1=0.9108635097493036,
best_f1=0.9154929577464789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045935477828606963
step: 10, loss: 5.398643406806514e-05
step: 20, loss: 7.43005657568574e-05
step: 30, loss: 9.871989459497854e-05
step: 40, loss: 0.00034768768819049
step: 50, loss: 0.000435155990999192
step: 60, loss: 4.6912140533095226e-05
step: 70, loss: 4.75190972792916e-05
step: 80, loss: 7.634317444171757e-05
step: 90, loss: 1.6368667274946347e-05
step: 100, loss: 4.269373675924726e-05
step: 110, loss: 0.015806078910827637
step: 120, loss: 0.0031230486929416656
step: 130, loss: 0.0010700931306928396
step: 140, loss: 7.369428203674033e-05
step: 150, loss: 0.000558774103410542
step: 160, loss: 8.538130350643769e-05
step: 170, loss: 7.176637882366776e-05
step: 180, loss: 0.0003561018966138363
step: 190, loss: 0.000172426865901798
step: 200, loss: 0.0019432323751971126
step: 210, loss: 0.030529074370861053
step: 220, loss: 8.191850793082267e-05
step: 230, loss: 0.00012844367302022874
step: 240, loss: 0.00014752522110939026
step: 250, loss: 0.0001357243600068614
step: 260, loss: 0.008419059216976166
step: 270, loss: 0.0002526802709326148
step: 280, loss: 0.0005964055308140814
step: 290, loss: 0.004635398741811514
step: 300, loss: 0.001587382983416319
step: 310, loss: 0.0044116126373410225
step: 320, loss: 0.0004872849676758051
step: 330, loss: 0.013307295739650726
step: 340, loss: 0.002253627637401223
step: 350, loss: 0.00019341045117471367
step: 360, loss: 0.0016679471591487527
step: 370, loss: 0.0005827161949127913
step: 380, loss: 7.531455048592761e-05
step: 390, loss: 4.558606815407984e-05
step: 400, loss: 0.0022228353191167116

```

```

step: 410, loss: 0.011387857608497143
step: 420, loss: 0.00024294573813676834
step: 430, loss: 3.864615428028628e-05
step: 440, loss: 0.06721518933773041
step: 450, loss: 0.0034310072660446167
step: 460, loss: 0.00017742265481501818
step: 470, loss: 0.00035252165980637074
step: 480, loss: 0.004762886092066765
step: 490, loss: 0.0007914243033155799
step: 500, loss: 0.03020257130265236
step: 510, loss: 0.0002941552666015923
step: 520, loss: 0.00010307473712600768
step: 530, loss: 3.78788827219978e-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=0.9159584513692162, f1=0.9123134328358209,
best_f1=0.9154929577464789
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 [00:16, 355.71it/s]
load_f1 = 0.9236499068901304
real_f1 = 0.9191449814126395
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 [00:12, 365.07it/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 distilbert \
  --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 distilbert \
  --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: 0.8697091937065125
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=0.2857142857142857,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35882970690727234

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4150943396226415, 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: 0.27853861451148987
/usr/local/lib/python3.7/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=0.38095238095238093, f1=0.42857142857142855,
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: 0.36071649193763733
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4680851063829786, f1=0.43478260869565216,
best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24034073948860168
/usr/local/lib/python3.7/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=0.45283018867924535, f1=0.44897959183673464,
best_f1=0.43478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2130606472492218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```


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=0.5161290322580646, f1=0.47058823529411764,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19815093278884888
/usr/local/lib/python3.7/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=0.5142857142857143, f1=0.3783783783783784,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28648748993873596
/usr/local/lib/python3.7/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=0.4363636363636364, f1=0.4705882352941177,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30643078684806824
/usr/local/lib/python3.7/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=0.4878048780487805, f1=0.4285714285714286,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25121164321899414
/usr/local/lib/python3.7/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=0.43750000000000006, f1=0.37931034482758624,
best_f1=0.47058823529411764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1922100931406021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5238095238095237, f1=0.4, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25010377168655396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5405405405405405, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12160138040781021
/usr/local/lib/python3.7/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=0.5405405405405405, f1=0.5000000000000001,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18045730888843536
/usr/local/lib/python3.7/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=0.5294117647058824, f1=0.5294117647058824,
best_f1=0.5000000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2701917290687561
/usr/local/lib/python3.7/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=0.5294117647058824, f1=0.5294117647058824,
best_f1=0.5000000000000001
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 125347.02it/s]
/usr/local/lib/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 = 0.5
real_f1 = 0.4905660377358491
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 [00:11, 391.21it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

```

```

--n_epochs 15 \
--finetuning \
--lm distilbert \
--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 distilbert \
  --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: 0.8130472302436829
step: 10, loss: 0.4812587797641754
step: 20, loss: 0.5728852152824402
step: 30, loss: 0.4671119451522827
step: 40, loss: 0.3480587899684906
step: 50, loss: 0.20742709934711456
step: 60, loss: 0.1327299177646637
step: 70, loss: 0.13604730367660522
step: 80, loss: 0.24467280507087708
step: 90, loss: 0.04467177391052246
step: 100, loss: 0.10032990574836731
step: 110, loss: 0.061194147914648056
step: 120, loss: 0.07058194279670715
step: 130, loss: 0.04306647554039955
step: 140, loss: 0.07223421335220337
step: 150, loss: 0.031111467629671097
step: 160, loss: 0.18410764634609222

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

step: 170, loss: 0.025991352275013924
step: 180, loss: 0.02062310464680195
step: 190, loss: 0.13085980713367462
step: 200, loss: 0.03520604968070984
step: 210, loss: 0.03930225968360901
step: 220, loss: 0.002881837310269475
step: 230, loss: 0.022085823118686676

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

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

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

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=0.9721913236929923, f1=0.9720044792833147,
best_f1=0.9720044792833147

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

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

step: 0, loss: 0.02623939700424671
step: 10, loss: 0.001009893137961626
step: 20, loss: 0.005512874107807875

```

step: 30, loss: 0.05959116294980049
step: 40, loss: 0.027110666036605835
step: 50, loss: 0.0048235259018838406
step: 60, loss: 0.004651535302400589
step: 70, loss: 0.015457064844667912
step: 80, loss: 0.1145830973982811
step: 90, loss: 0.028814353048801422
step: 100, loss: 0.013735183514654636
step: 110, loss: 0.005020813550800085
step: 120, loss: 0.0010805997299030423
step: 130, loss: 0.0025519058108329773
step: 140, loss: 0.14392225444316864
step: 150, loss: 0.008339395746588707
step: 160, loss: 0.00958988443017006
step: 170, loss: 0.09795772284269333
step: 180, loss: 0.006877618841826916
step: 190, loss: 0.1847635954618454
step: 200, loss: 0.10337311029434204
step: 210, loss: 0.03243541717529297
step: 220, loss: 0.0014981332933530211
step: 230, loss: 0.058047957718372345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808342728297633, 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: 0.034420039504766464
step: 10, loss: 0.05566832423210144
step: 20, loss: 0.011896872892975807
step: 30, loss: 0.0041204579174518585
step: 40, loss: 0.0021782175172120333
step: 50, loss: 0.0030775265768170357
step: 60, loss: 0.15163536369800568
step: 70, loss: 0.017042584717273712
step: 80, loss: 0.03639974445104599
step: 90, loss: 0.005557539872825146
step: 100, loss: 0.00701223686337471
step: 110, loss: 0.012505858205258846
step: 120, loss: 0.006333481520414352
step: 130, loss: 0.003808987094089389
step: 140, loss: 0.0012224860256537795

```

```

step: 150, loss: 0.021010875701904297
step: 160, loss: 0.16828498244285583
step: 170, loss: 0.0022331990767270327
step: 180, loss: 0.010182078927755356
step: 190, loss: 0.010452414862811565
step: 200, loss: 0.003884686855599284
step: 210, loss: 0.014309888705611229
step: 220, loss: 0.007485399022698402
step: 230, loss: 0.1637076735496521
/usr/local/lib/python3.7/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=0.98, f1=0.9571899012074643, 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: 0.0017565314192324877
step: 10, loss: 0.0024451082572340965
step: 20, loss: 0.009454591199755669
step: 30, loss: 0.0005531689384952188
step: 40, loss: 0.0016689551994204521
step: 50, loss: 0.060862354934215546
step: 60, loss: 0.003914743196219206
step: 70, loss: 0.002763211028650403
step: 80, loss: 0.012657318264245987
step: 90, loss: 0.0018936897395178676
step: 100, loss: 0.0014438880607485771
step: 110, loss: 0.005522988736629486
step: 120, loss: 0.005322127602994442
step: 130, loss: 0.0010110203875228763
step: 140, loss: 0.002891344018280506
step: 150, loss: 0.0011064667487517
step: 160, loss: 0.0017015421763062477
step: 170, loss: 0.007368545513600111
step: 180, loss: 0.03838327154517174
step: 190, loss: 0.004704958293586969
step: 200, loss: 0.001581378630362451
step: 210, loss: 0.0021669152192771435
step: 220, loss: 0.002107672393321991
step: 230, loss: 0.004181734286248684
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853438556933484, f1=0.9726651480637813,

```

```

best_f1=0.9726651480637813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014902657130733132
step: 10, loss: 0.0005303565412759781
step: 20, loss: 0.0002998769050464034
step: 30, loss: 0.00017746993398759514
step: 40, loss: 0.0005925337900407612
step: 50, loss: 0.0002696557203307748
step: 60, loss: 0.0003725184069480747
step: 70, loss: 0.0003021454031113535
step: 80, loss: 0.001585721387527883
step: 90, loss: 0.0022250560577958822
step: 100, loss: 0.00277496175840497
step: 110, loss: 0.0008889975142665207
step: 120, loss: 0.039368003606796265
step: 130, loss: 0.005967913195490837
step: 140, loss: 0.00022022411576472223
step: 150, loss: 0.00010992246825480834
step: 160, loss: 0.0009300938108935952
step: 170, loss: 0.0005377884372137487
step: 180, loss: 0.0011182299349457026
step: 190, loss: 0.00040904857451096177
step: 200, loss: 0.001697530155070126
step: 210, loss: 0.0001759974838932976
step: 220, loss: 0.00018542994803283364
step: 230, loss: 0.00036063871812075377
/usr/local/lib/python3.7/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=0.9733333333333333, f1=0.9753914988814317,
best_f1=0.9726651480637813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015429632039740682
step: 10, loss: 0.009217942133545876
step: 20, loss: 0.0003588232502806932
step: 30, loss: 0.0002851559256669134
step: 40, loss: 0.0007402607589028776
step: 50, loss: 0.0008952398202382028
step: 60, loss: 0.025322308763861656
step: 70, loss: 0.0001497894845670089
step: 80, loss: 0.0003758464299608022
step: 90, loss: 0.00012510899978224188
step: 100, loss: 0.00023603862791787833
step: 110, loss: 0.0747884213924408

```



```

step: 120, loss: 0.002629669150337577
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.15406019985675812
step: 140, loss: 0.0013915877789258957
step: 150, loss: 0.011330301873385906
step: 160, loss: 0.0022481651976704597
step: 170, loss: 0.0003212617593817413
step: 180, loss: 0.016498802229762077
step: 190, loss: 0.0033981616143137217
step: 200, loss: 0.012978574261069298
step: 210, loss: 0.0013119539944455028
step: 220, loss: 0.0005341452779248357
step: 230, loss: 0.000996425747871399
/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9773242630385486,
best_f1=0.9726651480637813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035964972339570522
step: 10, loss: 0.0004928666749037802
step: 20, loss: 0.0003376816166564822
step: 30, loss: 0.000669207307510078
step: 40, loss: 0.00015811446064617485
step: 50, loss: 0.0018048373749479651
step: 60, loss: 0.09365388751029968
step: 70, loss: 0.00048627820797264576
step: 80, loss: 0.007950253784656525
step: 90, loss: 0.00023544243595097214
step: 100, loss: 0.0006288365693762898
step: 110, loss: 0.0009359167306683958
step: 120, loss: 0.0008467150619253516
step: 130, loss: 0.09555703401565552
step: 140, loss: 0.001480644103139639
step: 150, loss: 0.0008004247792996466
step: 160, loss: 0.03300642967224121
step: 170, loss: 0.0016731400974094868
step: 180, loss: 0.005589915439486504
step: 190, loss: 0.0005514878430403769
step: 200, loss: 0.01023333240300417
step: 210, loss: 0.00021544347691815346
step: 220, loss: 0.00023032199533190578
step: 230, loss: 0.03711213544011116
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865168539325843, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03172982111573219
step: 10, loss: 0.0013724579475820065
step: 20, loss: 0.0013494695303961635
step: 30, loss: 0.0001576779905008152
step: 40, loss: 0.00012426823377609253
step: 50, loss: 0.0026077989023178816
step: 60, loss: 0.0010143055114895105
step: 70, loss: 0.00011059763346565887
step: 80, loss: 0.00027604150818660855
step: 90, loss: 0.0003064553311560303
step: 100, loss: 0.0005502264248207211
step: 110, loss: 0.0004315192636568099
step: 120, loss: 0.00016424106433987617
step: 130, loss: 0.03868795931339264
step: 140, loss: 0.00010692344221752137
step: 150, loss: 0.00021713192109018564
step: 160, loss: 0.00010258541442453861
step: 170, loss: 0.002105066552758217
step: 180, loss: 0.0017927956068888307
step: 190, loss: 0.0008193993126042187
step: 200, loss: 0.002728533698245883
step: 210, loss: 0.0014932004269212484
step: 220, loss: 0.001063019735738635
step: 230, loss: 0.0007411055848933756
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.9763779527559054,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007264355663210154
step: 10, loss: 0.009151764214038849
step: 20, loss: 0.00021892201039008796
step: 30, loss: 0.0032004464883357286
step: 40, loss: 8.244369382737204e-05
step: 50, loss: 0.0006353547214530408
step: 60, loss: 0.00016475444135721773

```

```

step: 70, loss: 0.0002867185103241354
step: 80, loss: 0.00040541531052440405
step: 90, loss: 0.0019266098970547318
step: 100, loss: 0.00023357008467428386
step: 110, loss: 0.00023754574067424983
step: 120, loss: 0.025763966143131256
step: 130, loss: 0.00013090237916912884
step: 140, loss: 0.0009121651528403163
step: 150, loss: 0.00010223268327536061
step: 160, loss: 0.0007613925845362246
step: 170, loss: 0.00014069629833102226
step: 180, loss: 0.00027022173162549734
step: 190, loss: 0.0001155144491349347
step: 200, loss: 0.00019519016495905817
step: 210, loss: 0.00011749160330509767
step: 220, loss: 0.049147602170705795
step: 230, loss: 0.0005531759234145284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9831271091113611, f1=0.9774774774774775,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003025334735866636
step: 10, loss: 5.950042759650387e-05
step: 20, loss: 0.0009367968305014074
step: 30, loss: 0.0007777531864121556
step: 40, loss: 0.00015260463987942785
step: 50, loss: 0.006344322115182877
step: 60, loss: 0.01172938384115696
step: 70, loss: 0.0006218077032826841
step: 80, loss: 0.0002753427252173424
step: 90, loss: 0.0002215040585724637
step: 100, loss: 0.00014084819122217596
step: 110, loss: 0.00011770464334404096
step: 120, loss: 5.627423888654448e-05
step: 130, loss: 9.693908941699192e-05
step: 140, loss: 0.13878217339515686
step: 150, loss: 0.00016640321700833738
step: 160, loss: 0.07309272885322571
step: 170, loss: 0.012876484543085098
step: 180, loss: 0.11018764972686768
step: 190, loss: 0.00017431701417081058
step: 200, loss: 0.00044573910417966545
step: 210, loss: 0.0001544394763186574
step: 220, loss: 0.0008620977168902755

```

```

step: 230, loss: 0.00067900464637205
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9751693002257337,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039978904533199966
step: 10, loss: 0.002535451902076602
step: 20, loss: 0.00011432635074015707
step: 30, loss: 0.000776634959038347
step: 40, loss: 0.00012367998715490103
step: 50, loss: 0.0003539695462677628
step: 60, loss: 0.00023235664411913604
step: 70, loss: 0.00024050108913797885
step: 80, loss: 0.03348312899470329
step: 90, loss: 0.0001375868741888553
step: 100, loss: 0.000583100481890142
step: 110, loss: 9.358003444503993e-05
step: 120, loss: 0.0006953741540201008
step: 130, loss: 0.00029674419783987105
step: 140, loss: 0.00013055272575002164
step: 150, loss: 0.00019337632693350315
step: 160, loss: 8.43654852360487e-05
step: 170, loss: 0.0004655623924918473
step: 180, loss: 0.0009312964975833893
step: 190, loss: 0.0023089395835995674
step: 200, loss: 0.0002465785073582083
step: 210, loss: 0.0007445589872077107
step: 220, loss: 0.00015000483836047351
step: 230, loss: 0.00024265662068501115
/usr/local/lib/python3.7/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=0.983277591973244, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.708469977136701e-05
step: 10, loss: 9.57679221755825e-05
step: 20, loss: 6.638823833782226e-05
step: 30, loss: 0.0012798603856936097
step: 40, loss: 9.933559340424836e-05
step: 50, loss: 8.256305591203272e-05
step: 60, loss: 0.0033626507502049208

```

```

step: 70, loss: 5.601426892098971e-05
step: 80, loss: 7.125001138774678e-05
step: 90, loss: 6.21099243289791e-05
step: 100, loss: 5.935138324275613e-05
step: 110, loss: 0.0003235597687307745
step: 120, loss: 0.00012415596575010568
step: 130, loss: 0.00014239642769098282
step: 140, loss: 0.00030180750763975084
step: 150, loss: 6.859091809019446e-05
step: 160, loss: 0.00020768755348399282
step: 170, loss: 0.00011901449033757672
step: 180, loss: 7.430324330925941e-05
step: 190, loss: 0.0004399714816827327
step: 200, loss: 0.0002664591884240508
step: 210, loss: 7.571302558062598e-05
step: 220, loss: 0.0005527646862901747
step: 230, loss: 0.00015640562924090773
/usr/local/lib/python3.7/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=0.9844097995545658, f1=0.972129319955407,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001403837522957474
step: 10, loss: 0.0006591082783415914
step: 20, loss: 6.44725005258806e-05
step: 30, loss: 0.00228568771854043
step: 40, loss: 9.630982094677165e-05
step: 50, loss: 5.192448588786647e-05
step: 60, loss: 0.0010626016883179545
step: 70, loss: 8.202652679756284e-05
step: 80, loss: 4.603505658451468e-05
step: 90, loss: 0.00010828395170392469
step: 100, loss: 9.745282295625657e-05
step: 110, loss: 4.1498318751109764e-05
step: 120, loss: 0.0006030744989402592
step: 130, loss: 0.00021066494809929281
step: 140, loss: 0.00040457939030602574
step: 150, loss: 0.019647367298603058
step: 160, loss: 5.8713016187539324e-05
step: 170, loss: 6.756205402780324e-05
step: 180, loss: 0.0001170904579339549
step: 190, loss: 0.00018189946422353387
step: 200, loss: 4.3167055991943926e-05
step: 210, loss: 0.0004542008391581476
step: 220, loss: 4.940916187479161e-05

```

```

step: 230, loss: 0.0002520166744943708
/usr/local/lib/python3.7/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=0.9830124575311437, f1=0.9738933030646991,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.85627741809003e-05
step: 10, loss: 5.0395323341945186e-05
step: 20, loss: 0.0004224607546348125
step: 30, loss: 9.425699681742117e-05
step: 40, loss: 5.672246697940864e-05
step: 50, loss: 0.00032548251328989863
step: 60, loss: 0.0005373609019443393
step: 70, loss: 0.00023415047326125205
step: 80, loss: 0.00014145976456347853
step: 90, loss: 0.00011862006795126945
step: 100, loss: 0.00010050906712422147
step: 110, loss: 0.0002606969210319221
step: 120, loss: 7.579296652693301e-05
step: 130, loss: 0.00013716949615627527
step: 140, loss: 0.00021076592383906245
step: 150, loss: 6.407988257706165e-05
step: 160, loss: 0.002654431154951453
step: 170, loss: 3.256944546592422e-05
step: 180, loss: 7.332323002628982e-05
step: 190, loss: 6.533348641823977e-05
step: 200, loss: 4.4891934521729127e-05
step: 210, loss: 0.0019156549824401736
step: 220, loss: 0.000119778786029201
step: 230, loss: 3.558653043000959e-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=0.9841986455981941, f1=0.9739524348810873,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.416755837155506e-05
step: 10, loss: 0.00020362182112876326
step: 20, loss: 3.959485547966324e-05
step: 30, loss: 7.89731930126436e-05
step: 40, loss: 0.0001030469429679215
step: 50, loss: 0.0001254420494660735
step: 60, loss: 0.00010842568735824898

```

```

step: 70, loss: 7.594385533593595e-05
step: 80, loss: 6.87455030856654e-05
step: 90, loss: 3.152646968374029e-05
step: 100, loss: 0.00018864752200897783
step: 110, loss: 8.22080546640791e-05
step: 120, loss: 6.0474572819657624e-05
step: 130, loss: 6.202771328389645e-05
step: 140, loss: 6.86158673488535e-05
step: 150, loss: 8.764373342273757e-05
step: 160, loss: 0.00011645507765933871
step: 170, loss: 0.00011146043834742159
step: 180, loss: 9.294918709201738e-05
step: 190, loss: 8.590662764618173e-05
step: 200, loss: 6.0246242355788127e-05
step: 210, loss: 0.00012934049300383776
step: 220, loss: 7.689095946261659e-05
step: 230, loss: 7.509737770305946e-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=0.9830890642615557, f1=0.9751131221719457,
best_f1=0.9764309764309763
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```
warnings.warn("An input tensor was not cuda.")
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 [00:07, 322.39it/s]
load_f1 = 0.9865168539325843
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 [00:11, 388.39it/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 distilbert \
--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 distilbert \
--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: 0.8012093901634216

step: 10, loss: 0.4234168231487274

step: 20, loss: 0.5041611790657043

step: 30, loss: 0.46065405011177063

step: 40, loss: 0.3819788694381714

step: 50, loss: 0.207436665892601

step: 60, loss: 0.17974939942359924

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: 70, loss: 0.14468863606452942

step: 80, loss: 0.11331794410943985

step: 90, loss: 0.16071346402168274

step: 100, loss: 0.2718469798564911

step: 110, loss: 0.1512860506772995

step: 120, loss: 0.10030650347471237

```

step: 130, loss: 0.028073862195014954
step: 140, loss: 0.2766556441783905
step: 150, loss: 0.02234373427927494
step: 160, loss: 0.12482677400112152
step: 170, loss: 0.2473519742488861
step: 180, loss: 0.10489866882562637
step: 190, loss: 0.10057540237903595
step: 200, loss: 0.12588950991630554
step: 210, loss: 0.11175438016653061
step: 220, loss: 0.10914228856563568
step: 230, loss: 0.1377740353345871
step: 240, loss: 0.13042278587818146
step: 250, loss: 0.1563670039176941
step: 260, loss: 0.07265172898769379
step: 270, loss: 0.029929788783192635
step: 280, loss: 0.22430576384067535
step: 290, loss: 0.03621162101626396
step: 300, loss: 0.084007129073143
step: 310, loss: 0.13849003612995148
step: 320, loss: 0.10305912047624588
step: 330, loss: 0.13374099135398865
step: 340, loss: 0.2132657766342163
step: 350, loss: 0.12025175988674164
step: 360, loss: 0.11921122670173645
step: 370, loss: 0.13528592884540558
step: 380, loss: 0.28856027126312256
step: 390, loss: 0.05687893182039261
step: 400, loss: 0.05215701088309288
step: 410, loss: 0.06990299373865128
step: 420, loss: 0.04882793873548508
step: 430, loss: 0.09435128420591354
step: 440, loss: 0.17364728450775146
step: 450, loss: 0.043788373470306396
step: 460, loss: 0.07445947825908661
step: 470, loss: 0.34001821279525757
step: 480, loss: 0.2088242471218109
step: 490, loss: 0.06013941019773483
step: 500, loss: 0.05893322080373764
step: 510, loss: 0.05521809682250023
step: 520, loss: 0.15339073538780212
step: 530, loss: 0.12298405170440674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9041970802919708, f1=0.9106654512306289,  
best_f1=0.9106654512306289  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.17325207591056824  
step: 10, loss: 0.1406451314687729  
step: 20, loss: 0.13173922896385193  
step: 30, loss: 0.05405816808342934  
step: 40, loss: 0.013729908503592014  
step: 50, loss: 0.06537917256355286  
step: 60, loss: 0.18665795028209686  
step: 70, loss: 0.1611817181110382  
step: 80, loss: 0.03381115198135376  
step: 90, loss: 0.058974843472242355  
step: 100, loss: 0.24135059118270874  
step: 110, loss: 0.03781435266137123  
step: 120, loss: 0.07962635904550552  
step: 130, loss: 0.08600421249866486  
step: 140, loss: 0.08107654750347137  
step: 150, loss: 0.07836107909679413  
step: 160, loss: 0.03422810509800911  
step: 170, loss: 0.10694074630737305  
step: 180, loss: 0.004369149450212717  
step: 190, loss: 0.07112761586904526  
step: 200, loss: 0.027022335678339005  
step: 210, loss: 0.05519406497478485  
step: 220, loss: 0.1264588087797165  
step: 230, loss: 0.17663417756557465  
step: 240, loss: 0.13510607182979584  
step: 250, loss: 0.13917435705661774  
step: 260, loss: 0.046329423785209656  
step: 270, loss: 0.2573273777961731  
step: 280, loss: 0.22312039136886597  
step: 290, loss: 0.11654507368803024  
step: 300, loss: 0.030641498044133186  
step: 310, loss: 0.05969229340553284  
step: 320, loss: 0.07215319573879242  
step: 330, loss: 0.035287804901599884  
step: 340, loss: 0.0066329604014754295  
step: 350, loss: 0.03161616995930672  
step: 360, loss: 0.04706780984997749  
step: 370, loss: 0.004201146308332682  
step: 380, loss: 0.08052810281515121  
step: 390, loss: 0.03523755073547363  
step: 400, loss: 0.1178879663348198  
step: 410, loss: 0.008898925967514515  
step: 420, loss: 0.11019482463598251
```

```

step: 430, loss: 0.04775277525186539
step: 440, loss: 0.05536883696913719
step: 450, loss: 0.012126564048230648
step: 460, loss: 0.08861905336380005
step: 470, loss: 0.129266619682312
step: 480, loss: 0.18178796768188477
step: 490, loss: 0.08817055076360703
step: 500, loss: 0.04689959064126015
step: 510, loss: 0.015173475258052349
step: 520, loss: 0.02861185558140278
step: 530, loss: 0.15372565388679504
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.915270018621974, f1=0.9190189726978251,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03780411556363106
step: 10, loss: 0.08871670067310333
step: 20, loss: 0.24481168389320374
step: 30, loss: 0.17995746433734894
step: 40, loss: 0.03012934885919094
step: 50, loss: 0.01952916756272316
step: 60, loss: 0.0206278208643198
step: 70, loss: 0.06269043684005737
step: 80, loss: 0.01689133420586586
step: 90, loss: 0.05526965484023094
step: 100, loss: 0.05065718665719032
step: 110, loss: 0.09667583554983139
step: 120, loss: 0.02397189661860466
step: 130, loss: 0.04212163761258125
step: 140, loss: 0.028708351776003838
step: 150, loss: 0.004348913207650185
step: 160, loss: 0.011541467159986496
step: 170, loss: 0.08273375034332275
step: 180, loss: 0.026539677754044533
step: 190, loss: 0.018278753384947777
step: 200, loss: 0.02558203600347042
step: 210, loss: 0.14524604380130768
step: 220, loss: 0.02301660180091858
step: 230, loss: 0.05621189996600151
step: 240, loss: 0.005879621021449566

```

```

step: 250, loss: 0.035819876939058304
step: 260, loss: 0.0027208521496504545
step: 270, loss: 0.004247238859534264
step: 280, loss: 0.007588213309645653
step: 290, loss: 0.04799935966730118
step: 300, loss: 0.03112965263426304
step: 310, loss: 0.020270247012376785
step: 320, loss: 0.1170227974653244
step: 330, loss: 0.01145889051258564
step: 340, loss: 0.010563489980995655
step: 350, loss: 0.019467001780867577
step: 360, loss: 0.0027178230229765177
step: 370, loss: 0.00789954699575901
step: 380, loss: 0.012103571556508541
step: 390, loss: 0.014615576714277267
step: 400, loss: 0.10052802413702011
step: 410, loss: 0.013025689870119095
step: 420, loss: 0.0293924268335104
step: 430, loss: 0.012818809598684311
step: 440, loss: 0.09092891961336136
step: 450, loss: 0.023777274414896965
step: 460, loss: 0.08742118626832962
step: 470, loss: 0.01988937333226204
step: 480, loss: 0.01247482281178236
step: 490, loss: 0.030619792640209198
step: 500, loss: 0.04594363644719124
step: 510, loss: 0.06078050658106804
step: 520, loss: 0.01918742060661316
step: 530, loss: 0.05397644266486168
/usr/local/lib/python3.7/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=0.9106067623899954, f1=0.9099307159353349,
best_f1=0.9190189726978251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008607410825788975
step: 10, loss: 0.13979007303714752
step: 20, loss: 0.04907393082976341
step: 30, loss: 0.09281772375106812
step: 40, loss: 0.0035901942756026983
step: 50, loss: 0.030107859522104263
step: 60, loss: 0.025012044236063957
step: 70, loss: 0.000474155560368672
step: 80, loss: 0.030629660934209824
step: 90, loss: 0.07853326201438904
step: 100, loss: 0.14308449625968933

```

```
step: 110, loss: 0.004699596669524908
step: 120, loss: 0.004937030375003815
step: 130, loss: 0.004697063006460667
step: 140, loss: 0.007459399290382862
step: 150, loss: 0.0011987951584160328
step: 160, loss: 0.003105888608843088
step: 170, loss: 0.0052333613857626915
step: 180, loss: 0.00986463762819767
step: 190, loss: 0.0015866104513406754
step: 200, loss: 0.0031651072204113007
step: 210, loss: 0.017402704805135727
step: 220, loss: 0.019746560603380203
step: 230, loss: 0.2031487077474594
step: 240, loss: 0.07123100757598877
step: 250, loss: 0.010500050149857998
step: 260, loss: 0.10589428246021271
step: 270, loss: 0.050046708434820175
step: 280, loss: 0.008013584651052952
step: 290, loss: 0.17333462834358215
step: 300, loss: 0.03208821639418602
step: 310, loss: 0.046860117465257645
step: 320, loss: 0.10281617939472198
step: 330, loss: 0.13849672675132751
step: 340, loss: 0.1632099598646164
step: 350, loss: 0.07272282987833023
step: 360, loss: 0.014806930907070637
step: 370, loss: 0.008061662316322327
step: 380, loss: 0.005898674018681049
step: 390, loss: 0.01929284632205963
step: 400, loss: 0.059213217347860336
step: 410, loss: 0.006373492535203695
step: 420, loss: 0.026096155866980553
step: 430, loss: 0.012858649715781212
step: 440, loss: 0.010288471356034279
step: 450, loss: 0.006778178736567497
step: 460, loss: 0.005939434748142958
step: 470, loss: 0.0016278899274766445
step: 480, loss: 0.027666360139846802
step: 490, loss: 0.012257687747478485
step: 500, loss: 0.00486388523131609
step: 510, loss: 0.010129684582352638
step: 520, loss: 0.08380507677793503
step: 530, loss: 0.0013871050905436277
/usr/local/lib/python3.7/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=0.911508482347547, f1=0.9176578225068619,
best_f1=0.9190189726978251
```

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

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

```
step: 0, loss: 0.006205862853676081  
step: 10, loss: 0.002401556121185422  
step: 20, loss: 0.005027119070291519  
step: 30, loss: 0.002146964194253087  
step: 40, loss: 0.005559730809181929  
step: 50, loss: 0.001306306105107069  
step: 60, loss: 0.002713714726269245  
step: 70, loss: 0.0017518197419121861  
step: 80, loss: 0.002021884312853217  
step: 90, loss: 0.0058355494402348995  
step: 100, loss: 0.0002990987850353122  
step: 110, loss: 0.00036997153074480593  
step: 120, loss: 0.0033152743708342314  
step: 130, loss: 0.0014961028937250376  
step: 140, loss: 0.0009944629855453968  
step: 150, loss: 0.002439424628391862  
step: 160, loss: 0.06418415904045105  
step: 170, loss: 0.017651179805397987  
step: 180, loss: 0.0009751132456585765  
step: 190, loss: 0.0005725874798372388  
step: 200, loss: 0.01682044006884098  
step: 210, loss: 0.002151777967810631  
step: 220, loss: 0.0004706434556283057  
step: 230, loss: 0.0017874061595648527  
step: 240, loss: 0.006056572310626507  
step: 250, loss: 0.000598046462982893  
step: 260, loss: 0.02604733034968376  
step: 270, loss: 0.1118941605091095  
step: 280, loss: 0.005417033098638058  
step: 290, loss: 0.0024317624047398567  
step: 300, loss: 0.022202756255865097  
step: 310, loss: 0.00026996078668162227  
step: 320, loss: 0.04787423461675644  
step: 330, loss: 0.0016830370295792818  
step: 340, loss: 0.0005849316949024796  
step: 350, loss: 0.0006189091363921762  
step: 360, loss: 0.05088207498192787  
step: 370, loss: 0.0008456227369606495  
step: 380, loss: 0.11296562850475311  
step: 390, loss: 0.039588067680597305  
step: 400, loss: 0.01528877206146717  
step: 410, loss: 0.00017849590221885592  
step: 420, loss: 0.004831859841942787  
step: 430, loss: 0.00574085908010602  
step: 440, loss: 0.0019301071297377348
```

```

step: 450, loss: 0.004040164407342672
step: 460, loss: 0.006157353054732084
step: 470, loss: 0.0021469308994710445
step: 480, loss: 0.004363223910331726
step: 490, loss: 0.037095699459314346
step: 500, loss: 0.05040794238448143
step: 510, loss: 0.006021668203175068
step: 520, loss: 0.0016563747776672244
step: 530, loss: 0.02053847536444664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9189435336976319, f1=0.9203459262630859,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008908200543373823
step: 10, loss: 0.042997512966394424
step: 20, loss: 0.0004255514359101653
step: 30, loss: 0.00658131530508399
step: 40, loss: 0.003752827411517501
step: 50, loss: 0.0012738100485876203
step: 60, loss: 0.01899641379714012
step: 70, loss: 0.02964920550584793
step: 80, loss: 0.0008912608027458191
step: 90, loss: 0.004717162810266018
step: 100, loss: 0.14293619990348816
step: 110, loss: 0.008485869504511356
step: 120, loss: 0.036484088748693466
step: 130, loss: 0.04256834089756012
step: 140, loss: 0.0069023300893604755
step: 150, loss: 0.0008689591777510941
step: 160, loss: 0.021880928426980972
step: 170, loss: 0.00034755244269035757
step: 180, loss: 0.000534032064024359
step: 190, loss: 0.004688342101871967
step: 200, loss: 0.0003669751458801329
step: 210, loss: 0.0005563675658777356
step: 220, loss: 0.00021662842482328415
step: 230, loss: 0.0007348655490204692
step: 240, loss: 0.0018682221416383982
step: 250, loss: 0.00041655541281215847
step: 260, loss: 0.003204277018085122

```



```

step: 270, loss: 0.0032253284007310867
step: 280, loss: 0.002344476757571101
step: 290, loss: 0.0006947172223590314
step: 300, loss: 0.057775989174842834
step: 310, loss: 0.008950017392635345
step: 320, loss: 0.03372519090771675
step: 330, loss: 0.0031485133804380894
step: 340, loss: 0.0008251882973127067
step: 350, loss: 0.02870582975447178
step: 360, loss: 0.0007951518637128174
step: 370, loss: 0.007865115068852901
step: 380, loss: 0.011191771365702152
step: 390, loss: 0.09126491099596024
step: 400, loss: 0.00022080694907344878
step: 410, loss: 0.0005465778522193432
step: 420, loss: 0.10127867758274078
step: 430, loss: 0.005827941000461578
step: 440, loss: 0.0014114804798737168
step: 450, loss: 0.0021451429929584265
step: 460, loss: 0.0006344428984448314
step: 470, loss: 0.0010144665138795972
step: 480, loss: 0.004246963188052177
step: 490, loss: 0.00015532298129983246
step: 500, loss: 0.004034551326185465
step: 510, loss: 0.00029421798535622656
step: 520, loss: 0.0005420717061497271
step: 530, loss: 0.0006314925849437714
/usr/local/lib/python3.7/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=0.8927727916863487, f1=0.89185326345879,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025026097893714905
step: 10, loss: 0.08291167765855789
step: 20, loss: 0.004791250452399254
step: 30, loss: 0.000653265044093132
step: 40, loss: 6.483621837105602e-05
step: 50, loss: 0.0005326270475052297
step: 60, loss: 0.0004245895834174007
step: 70, loss: 0.0005307610845193267
step: 80, loss: 0.010025176219642162
step: 90, loss: 0.014465566724538803
step: 100, loss: 0.0007674708031117916
step: 110, loss: 0.0003807775501627475
step: 120, loss: 0.00028083508368581533

```

```

step: 130, loss: 0.000295670994091779
step: 140, loss: 0.001128257717937231
step: 150, loss: 9.235045581590384e-05
step: 160, loss: 0.0004372127295937389
step: 170, loss: 0.0011852015741169453
step: 180, loss: 0.03922230005264282
step: 190, loss: 0.009895890951156616
step: 200, loss: 0.0015166030498221517
step: 210, loss: 0.00033417262602597475
step: 220, loss: 0.0009600131306797266
step: 230, loss: 0.0005414137267507613
step: 240, loss: 0.0024714109022170305
step: 250, loss: 0.0011886288411915302
step: 260, loss: 0.0043206000700592995
step: 270, loss: 0.00021413486683741212
step: 280, loss: 0.01895091123878956
step: 290, loss: 0.0008367444388568401
step: 300, loss: 0.00036547830677591264
step: 310, loss: 0.00028148002456873655
step: 320, loss: 0.07507704198360443
step: 330, loss: 6.0912112530786544e-05
step: 340, loss: 0.021588794887065887
step: 350, loss: 0.016343407332897186
step: 360, loss: 0.0020699393935501575
step: 370, loss: 0.0006741562974639237
step: 380, loss: 0.00015604848158545792
step: 390, loss: 9.99513067654334e-05
step: 400, loss: 7.983767136465758e-05
step: 410, loss: 0.000308756047161296
step: 420, loss: 0.0005194597179070115
step: 430, loss: 5.249754758551717e-05
step: 440, loss: 0.013448012992739677
step: 450, loss: 0.005926310550421476
step: 460, loss: 0.018089136108756065
step: 470, loss: 0.044128332287073135
step: 480, loss: 0.11142922937870026
step: 490, loss: 0.04250925034284592
step: 500, loss: 0.00045920288539491594
step: 510, loss: 0.0008398817735724151
step: 520, loss: 0.0019925080705434084
step: 530, loss: 0.0028383799362927675
/usr/local/lib/python3.7/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=0.9132250580046404, f1=0.9138888888888889,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021878553088754416
step: 10, loss: 0.0016036295564845204
step: 20, loss: 0.006148139480501413
step: 30, loss: 0.001990938326343894
step: 40, loss: 0.0009456003317609429
step: 50, loss: 0.0003235302574466914
step: 60, loss: 0.00025180590455420315
step: 70, loss: 0.06062901020050049
step: 80, loss: 0.0004596231447067112
step: 90, loss: 0.0002153554087271914
step: 100, loss: 0.004273136146366596
step: 110, loss: 0.02724630758166313
step: 120, loss: 0.002615123288705945
step: 130, loss: 0.0003235600597690791
step: 140, loss: 5.0370807002764195e-05
step: 150, loss: 0.0002773019950836897
step: 160, loss: 0.0007900393102318048
step: 170, loss: 0.0005981158465147018
step: 180, loss: 0.0019940738566219807
step: 190, loss: 0.0001479113125242293
step: 200, loss: 0.00021070010552648455
step: 210, loss: 0.0025258588138967752
step: 220, loss: 0.0023936943616718054
step: 230, loss: 0.0075884005054831505
step: 240, loss: 0.00022346555488184094
step: 250, loss: 0.0019634338095784187
step: 260, loss: 0.02481430023908615
step: 270, loss: 5.132332080393098e-05
step: 280, loss: 0.00011866174463648349
step: 290, loss: 0.00032486827694810927
step: 300, loss: 0.00014731018745806068
step: 310, loss: 0.0024094555992633104
step: 320, loss: 0.000997127266600728
step: 330, loss: 0.0006845039897598326
step: 340, loss: 0.002096885349601507
step: 350, loss: 0.00033567374339327216
step: 360, loss: 0.00014982417633291334
step: 370, loss: 0.00012463616440072656
step: 380, loss: 0.04188242927193642
step: 390, loss: 0.006057612132281065
step: 400, loss: 0.20530818402767181
step: 410, loss: 0.005080599803477526
step: 420, loss: 0.14205549657344818
step: 430, loss: 0.0005126773612573743
step: 440, loss: 0.003727838397026062
step: 450, loss: 0.0004282125155441463
step: 460, loss: 0.0028182587120682
```

```

step: 470, loss: 0.0002663369523361325
step: 480, loss: 0.00041424643131904304
step: 490, loss: 0.0007762305322103202
step: 500, loss: 0.001348277204670012
step: 510, loss: 0.00021236117754597217
step: 520, loss: 0.0002288001705892384
step: 530, loss: 0.00012539837916847318
/usr/local/lib/python3.7/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=0.9042110134197131, f1=0.9074766355140187,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009246030240319669
step: 10, loss: 0.000249282835284248
step: 20, loss: 0.0003067799552809447
step: 30, loss: 0.0012747392756864429
step: 40, loss: 0.0030162385664880276
step: 50, loss: 0.0005655891145579517
step: 60, loss: 0.00011277126759523526
step: 70, loss: 0.027717778459191322
step: 80, loss: 0.0001606446603545919
step: 90, loss: 0.00014852908498141915
step: 100, loss: 0.00022110564168542624
step: 110, loss: 0.0015447060577571392
step: 120, loss: 0.0009627376566641033
step: 130, loss: 0.0008048602030612528
step: 140, loss: 0.04298939183354378
step: 150, loss: 0.0012484982144087553
step: 160, loss: 0.0014870916493237019
step: 170, loss: 9.030148794408888e-05
step: 180, loss: 9.156779560726136e-05
step: 190, loss: 0.004106045234948397
step: 200, loss: 0.0008264561765827239
step: 210, loss: 0.00013832167314831167
step: 220, loss: 0.001191899529658258
step: 230, loss: 0.00010895127343246713
step: 240, loss: 0.06303804367780685
step: 250, loss: 0.00026581023121252656
step: 260, loss: 0.00010530793952057138
step: 270, loss: 0.0006752845947630703
step: 280, loss: 8.363622328033671e-05
step: 290, loss: 6.1650644056499e-05
step: 300, loss: 6.420657882699743e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 310, loss: 0.0001675745443208143

```

```

step: 320, loss: 0.0013942092191427946
step: 330, loss: 0.00017594208475202322
step: 340, loss: 0.0001887945836642757
step: 350, loss: 0.0003186306275893003
step: 360, loss: 0.0012806103331968188
step: 370, loss: 0.00013589696027338505
step: 380, loss: 0.00034245155984535813
step: 390, loss: 5.879026866750792e-05
step: 400, loss: 0.015689866617321968
step: 410, loss: 0.0005402560927905142
step: 420, loss: 0.00022383687610272318
step: 430, loss: 0.0013413232518360019
step: 440, loss: 0.0002570466895122081
step: 450, loss: 0.00010717179247876629
step: 460, loss: 0.002861924469470978
step: 470, loss: 0.00016221099940594286
step: 480, loss: 8.685907232575119e-05
step: 490, loss: 7.507095870096236e-05
step: 500, loss: 0.00020547329040709883
step: 510, loss: 0.0006235046312212944
step: 520, loss: 0.00010763792670331895
step: 530, loss: 0.0017638576682657003
/usr/local/lib/python3.7/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=0.9110189027201476, f1=0.9186691312384473,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005580889410339296
step: 10, loss: 5.007685467717238e-05
step: 20, loss: 0.00025831005768850446
step: 30, loss: 0.00018105479830410331
step: 40, loss: 0.001309118582867086
step: 50, loss: 0.00012731696187984198
step: 60, loss: 0.00023466581478714943
step: 70, loss: 0.004132562316954136
step: 80, loss: 0.0012142866617068648
step: 90, loss: 0.0003007237974088639
step: 100, loss: 0.0010735688265413046
step: 110, loss: 0.0008299005567096174
step: 120, loss: 0.0013253476936370134
step: 130, loss: 0.0004806034849025309
step: 140, loss: 0.00031402637250721455
step: 150, loss: 0.00012354318459983915
step: 160, loss: 0.00010949109127977863
step: 170, loss: 5.316478927852586e-05

```

```

step: 180, loss: 0.0003188482951372862
step: 190, loss: 0.0001483826490584761
step: 200, loss: 9.551036782795563e-05
step: 210, loss: 7.242038554977626e-05
step: 220, loss: 0.00025269901379942894
step: 230, loss: 4.6727571316296235e-05
step: 240, loss: 6.980547914281487e-05
step: 250, loss: 0.0005219620652496815
step: 260, loss: 0.0011244503548368812
step: 270, loss: 5.547456385102123e-05
step: 280, loss: 0.0004162836412433535
step: 290, loss: 0.0003168589319102466
step: 300, loss: 0.013067949563264847
step: 310, loss: 5.6857727031456307e-05
step: 320, loss: 4.628744500223547e-05
step: 330, loss: 0.00012063998292433098
step: 340, loss: 5.854669871041551e-05
step: 350, loss: 0.00010303977614967152
step: 360, loss: 7.8325487265829e-05
step: 370, loss: 0.11962512135505676
step: 380, loss: 5.3890416893409565e-05
step: 390, loss: 7.931320578791201e-05
step: 400, loss: 0.0001307497441302985
step: 410, loss: 0.00028042722260579467
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 420, loss: 0.00010916151222772896
step: 430, loss: 5.7946013839682564e-05
step: 440, loss: 8.163535676430911e-05
step: 450, loss: 0.0005869590677320957
step: 460, loss: 0.0006312831537798047
step: 470, loss: 0.00028870775713585317
step: 480, loss: 0.00018739697406999767
step: 490, loss: 0.008670180104672909
step: 500, loss: 0.051778748631477356
step: 510, loss: 0.00016293275984935462
step: 520, loss: 0.15899448096752167
step: 530, loss: 0.0019938249606639147
/usr/local/lib/python3.7/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=0.9064814814814813, f1=0.9091755938518863,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029403220396488905
step: 10, loss: 0.005236916244029999
step: 20, loss: 0.00010582675167825073

```

step: 30, loss: 0.04768781736493111
step: 40, loss: 5.1149443606846035e-05
step: 50, loss: 0.00037446891656145453
step: 60, loss: 0.0004301267908886075
step: 70, loss: 9.339149983134121e-05
step: 80, loss: 0.000317270023515448
step: 90, loss: 0.00013120254152454436
step: 100, loss: 0.00011506034206831828
step: 110, loss: 5.781337677035481e-05
step: 120, loss: 0.00020738309831358492
step: 130, loss: 0.00011529143375810236
step: 140, loss: 4.745454134535976e-05
step: 150, loss: 7.26161160855554e-05
step: 160, loss: 0.00011297575838398188
step: 170, loss: 0.0017504551215097308
step: 180, loss: 4.9115515139419585e-05
step: 190, loss: 0.00011455138883320615
step: 200, loss: 0.00012262718519195914
step: 210, loss: 9.06862915144302e-05
step: 220, loss: 0.00018844807345885783
step: 230, loss: 5.642740870825946e-05
step: 240, loss: 0.00022633877233602107
step: 250, loss: 0.00749779911711812
step: 260, loss: 0.00015539197192993015
step: 270, loss: 0.0003709853917825967
step: 280, loss: 0.0014132626820355654
step: 290, loss: 0.0006569977849721909
step: 300, loss: 0.004925239831209183
step: 310, loss: 0.00046369381016120315
step: 320, loss: 0.00450053159147501
step: 330, loss: 0.0005403810064308345
step: 340, loss: 0.047626182436943054
step: 350, loss: 0.002869653282687068
step: 360, loss: 0.0004695333482231945
step: 370, loss: 6.458916323026642e-05
step: 380, loss: 0.0002942463324870914
step: 390, loss: 0.006084281485527754
step: 400, loss: 4.388841989566572e-05
step: 410, loss: 0.004141312092542648
step: 420, loss: 0.00029576633824035525
step: 430, loss: 6.815759115852416e-05
step: 440, loss: 0.0008540365379303694
step: 450, loss: 0.0010587609140202403
step: 460, loss: 0.00015406844613607973
step: 470, loss: 0.0002665814245119691
step: 480, loss: 4.2321549699408934e-05
step: 490, loss: 0.0017040297389030457
step: 500, loss: 6.515155837405473e-05

```

step: 510, loss: 5.553643131861463e-05
step: 520, loss: 3.613794251577929e-05
step: 530, loss: 0.00020554526417981833
/usr/local/lib/python3.7/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=0.9095992544268405, f1=0.9124293785310734,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.816022844053805e-05
step: 10, loss: 9.392333595314994e-05
step: 20, loss: 5.060131297796033e-05
step: 30, loss: 4.5562217565020546e-05
step: 40, loss: 0.0001653897197684273
step: 50, loss: 4.604574860422872e-05
step: 60, loss: 5.772713484475389e-05
step: 70, loss: 0.00013272833894006908
step: 80, loss: 0.0005882055847905576
step: 90, loss: 0.0002580667787697166
step: 100, loss: 0.0003280662640463561
step: 110, loss: 4.300979708204977e-05
step: 120, loss: 5.38150779902935e-05
step: 130, loss: 0.00035719870356842875
step: 140, loss: 6.256144843064249e-05
step: 150, loss: 8.87073329067789e-05
step: 160, loss: 0.0001454939047107473
step: 170, loss: 5.281272024149075e-05
step: 180, loss: 3.710266901180148e-05
step: 190, loss: 5.088952093501575e-05
step: 200, loss: 4.499168790061958e-05
step: 210, loss: 0.0001855302252806723
step: 220, loss: 0.00021245362586341798
step: 230, loss: 0.004965738393366337
step: 240, loss: 5.06901960761752e-05
step: 250, loss: 0.00021527970966417342
step: 260, loss: 6.121164915384725e-05
step: 270, loss: 8.150414214469492e-05
step: 280, loss: 0.0030408373568207026
step: 290, loss: 0.0001582180120749399
step: 300, loss: 8.732340211281553e-05
step: 310, loss: 0.00011272887059021741
step: 320, loss: 4.007851384812966e-05
step: 330, loss: 0.00018319454102311283
step: 340, loss: 2.8706290322588757e-05
step: 350, loss: 0.001542069367133081
step: 360, loss: 0.00022633728804066777

```



```

step: 370, loss: 0.00011213105608476326
step: 380, loss: 6.882043089717627e-05
step: 390, loss: 4.21273653046228e-05
step: 400, loss: 0.000253558624535799
step: 410, loss: 6.25858738203533e-05
step: 420, loss: 0.00037050448008812964
step: 430, loss: 0.00011568387708393857
step: 440, loss: 0.00018402036221232265
step: 450, loss: 6.198508344823495e-05
step: 460, loss: 0.00010871696576941758
step: 470, loss: 0.00011995080421911553
step: 480, loss: 6.421087891794741e-05
step: 490, loss: 6.198144546942785e-05
step: 500, loss: 0.0024934043176472187
step: 510, loss: 0.0021362323313951492
step: 520, loss: 0.00017270416719838977
step: 530, loss: 0.00016461810446344316
/usr/local/lib/python3.7/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=0.9071494893221911, f1=0.9126576366184025,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.231169102946296e-05
step: 10, loss: 9.622705692891032e-05
step: 20, loss: 9.640750795369968e-05
step: 30, loss: 0.00010201707482337952
step: 40, loss: 3.843962258542888e-05
step: 50, loss: 6.753173511242494e-05
step: 60, loss: 3.551548070390709e-05
step: 70, loss: 3.1153304007602856e-05
step: 80, loss: 4.217798777972348e-05
step: 90, loss: 0.0005167878116481006
step: 100, loss: 3.4970846172655e-05
step: 110, loss: 4.399081808514893e-05
step: 120, loss: 3.9690141420578584e-05
step: 130, loss: 0.002667102264240384
step: 140, loss: 0.0002277588937431574
step: 150, loss: 7.062111399136484e-05
step: 160, loss: 3.473723336355761e-05
step: 170, loss: 3.100463800365105e-05
step: 180, loss: 0.0001241002610186115
step: 190, loss: 5.9961628721794114e-05
step: 200, loss: 2.5338797058793716e-05
step: 210, loss: 0.0005371986771933734
step: 220, loss: 3.5742759791901335e-05

```

```

step: 230, loss: 2.0671319362008944e-05
step: 240, loss: 0.000516225234605372
step: 250, loss: 0.004995351657271385
step: 260, loss: 3.680030204122886e-05
step: 270, loss: 3.12392512569204e-05
step: 280, loss: 4.820874528377317e-05
step: 290, loss: 0.0001531639281893149
step: 300, loss: 4.7908433771226555e-05
step: 310, loss: 0.038793954998254776
step: 320, loss: 4.239785630488768e-05
step: 330, loss: 8.688664092915133e-05
step: 340, loss: 3.5656896216096357e-05
step: 350, loss: 0.00033563782926648855
step: 360, loss: 3.037488932022825e-05
step: 370, loss: 2.1718056814279407e-05
step: 380, loss: 0.01940184086561203
step: 390, loss: 2.7622350899036974e-05
step: 400, loss: 3.617943366407417e-05
step: 410, loss: 3.4278924431419e-05
step: 420, loss: 5.1302591600688174e-05
step: 430, loss: 0.035581208765506744
step: 440, loss: 8.075062214629725e-05
step: 450, loss: 3.170865238644183e-05
step: 460, loss: 2.3159516786108725e-05
step: 470, loss: 4.252096914569847e-05
step: 480, loss: 3.241159356548451e-05
step: 490, loss: 0.0001276372786378488
step: 500, loss: 9.747493459144607e-05
step: 510, loss: 4.856879968428984e-05
step: 520, loss: 2.8445394491427578e-05
step: 530, loss: 7.842257764423266e-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=0.9119409866297833, f1=0.916044776119403,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039573493995703757
step: 10, loss: 0.0014256472932174802
step: 20, loss: 3.399825800443068e-05
step: 30, loss: 0.000987361534498632
step: 40, loss: 6.000701250741258e-05
step: 50, loss: 0.00038339581806212664
step: 60, loss: 3.9616435969946906e-05
step: 70, loss: 2.436280919937417e-05
step: 80, loss: 3.3820524549810216e-05

```

step: 90, loss: 1.9650486137834378e-05
step: 100, loss: 0.00020168286573607475
step: 110, loss: 2.5148696295218542e-05
step: 120, loss: 9.249684808310121e-05
step: 130, loss: 5.400466397986747e-05
step: 140, loss: 0.005456568207591772
step: 150, loss: 5.1931019697804004e-05
step: 160, loss: 2.5089178961934522e-05
step: 170, loss: 4.9080452299676836e-05
step: 180, loss: 2.5874942366499454e-05
step: 190, loss: 5.684175994247198e-05
step: 200, loss: 2.722748467931524e-05
step: 210, loss: 0.00015303457621484995
step: 220, loss: 2.2522666768054478e-05
step: 230, loss: 0.00011982655269093812
step: 240, loss: 3.099309469689615e-05
step: 250, loss: 6.621264765271917e-05
step: 260, loss: 1.661456553847529e-05
step: 270, loss: 8.906287257559597e-05
step: 280, loss: 3.020667099917773e-05
step: 290, loss: 2.7055821192334406e-05
step: 300, loss: 3.070530146942474e-05
step: 310, loss: 0.0003798907855525613
step: 320, loss: 3.9225142245413736e-05
step: 330, loss: 3.110092075075954e-05
step: 340, loss: 6.084766937419772e-05
step: 350, loss: 3.0077144401730038e-05
step: 360, loss: 4.6689154260093346e-05
step: 370, loss: 5.0349743105471134e-05
step: 380, loss: 2.3852400772739202e-05
step: 390, loss: 3.2708285289118066e-05
step: 400, loss: 3.7247904401738197e-05
step: 410, loss: 3.230068978155032e-05
step: 420, loss: 0.0001737335987854749
step: 430, loss: 2.748055885604117e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 440, loss: 1.712858102109749e-05
step: 450, loss: 2.6739438908407465e-05
step: 460, loss: 3.1671141186961904e-05
step: 470, loss: 2.747293547145091e-05
step: 480, loss: 2.797957495204173e-05
step: 490, loss: 3.071229366469197e-05
step: 500, loss: 6.20523642282933e-05
step: 510, loss: 6.124167703092098e-05
step: 520, loss: 4.849749893764965e-05
step: 530, loss: 0.00012374013022053987
/usr/local/lib/python3.7/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=0.9124423963133641, f1=0.9182915506035284,
best_f1=0.9203459262630859
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.8756487406790257e-05
step: 10, loss: 3.378324254299514e-05
step: 20, loss: 2.6933128538075835e-05
step: 30, loss: 0.00010019902401836589
step: 40, loss: 2.1401272533694282e-05
step: 50, loss: 2.067481545964256e-05
step: 60, loss: 6.104661588324234e-05
step: 70, loss: 4.2057446989929304e-05
step: 80, loss: 2.9234668545541354e-05
step: 90, loss: 2.2533637093147263e-05
step: 100, loss: 3.3789601729949936e-05
step: 110, loss: 0.00018407490279059857
step: 120, loss: 5.3557065257336944e-05
step: 130, loss: 2.4596878574811853e-05
step: 140, loss: 2.8560303690028377e-05
step: 150, loss: 4.0712489862926304e-05
step: 160, loss: 7.638891111128032e-05
step: 170, loss: 3.1264622521121055e-05
step: 180, loss: 7.533892494393513e-05
step: 190, loss: 8.073745993897319e-05
step: 200, loss: 2.589353061921429e-05
step: 210, loss: 0.009537270292639732
step: 220, loss: 1.4342199392558541e-05
step: 230, loss: 2.5394641852471977e-05
step: 240, loss: 4.21686090703588e-05
step: 250, loss: 9.438549750484526e-05
step: 260, loss: 2.4947512429207563e-05
step: 270, loss: 3.522501719999127e-05
step: 280, loss: 3.4353372029727325e-05
step: 290, loss: 0.00010783912148326635
step: 300, loss: 0.0011295481817796826
step: 310, loss: 0.00018047165940515697
step: 320, loss: 2.6758014428196475e-05
step: 330, loss: 0.00015144622011575848
step: 340, loss: 6.017976920702495e-05
step: 350, loss: 2.8467293304856867e-05
step: 360, loss: 6.11498107900843e-05
step: 370, loss: 0.00014123489381745458
step: 380, loss: 1.8138112864107825e-05
step: 390, loss: 0.007326503749936819
step: 400, loss: 2.172527092625387e-05
step: 410, loss: 8.00181514932774e-05

```

```

step: 420, loss: 0.00012787131709046662
step: 430, loss: 3.372271385160275e-05
step: 440, loss: 3.9113689126679674e-05
step: 450, loss: 4.6837714762659743e-05
step: 460, loss: 0.0004537613713182509
step: 470, loss: 3.191599535057321e-05
step: 480, loss: 0.0031506922096014023
step: 490, loss: 2.9548133170465007e-05
step: 500, loss: 3.6263736546970904e-05
step: 510, loss: 2.7108006179332733e-05
step: 520, loss: 2.638162004586775e-05
step: 530, loss: 2.6042522222269326e-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=0.9116022099447515, f1=0.9208699676075892,
best_f1=0.9203459262630859
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 [00:14, 387.82it/s]

```

```
load_f1 = 0.9186733303044072
real_f1 = 0.9193840579710145
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 [00:11, 395.89it/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 distilbert \
  --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 distilbert \
  --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: 0.8452023863792419
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.07113571465015411
step: 20, loss: 0.37670543789863586
step: 30, loss: 0.3591206669807434
step: 40, loss: 0.515080988407135
step: 50, loss: 0.31462788581848145
step: 60, loss: 0.38506045937538147
step: 70, loss: 0.22964398562908173
step: 80, loss: 0.2848650813102722
step: 90, loss: 0.41018712520599365
step: 100, loss: 0.1805475503206253
step: 110, loss: 0.3693223297595978
step: 120, loss: 0.2662472128868103
step: 130, loss: 0.2527722120285034
step: 140, loss: 0.2554102838039398
step: 150, loss: 0.3118913471698761
step: 160, loss: 0.2304413914680481
step: 170, loss: 0.18735647201538086
step: 180, loss: 0.20286330580711365
step: 190, loss: 0.23968636989593506

```

```

step: 200, loss: 0.17288482189178467
step: 210, loss: 0.4141726493835449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5162601626016261, f1=0.5152838427947598,
best_f1=0.5152838427947598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10663363337516785
step: 10, loss: 0.11321589350700378
step: 20, loss: 0.27248936891555786
step: 30, loss: 0.26624342799186707
step: 40, loss: 0.1205027624964714
step: 50, loss: 0.16045798361301422
step: 60, loss: 0.1084064170718193
step: 70, loss: 0.23029424250125885
step: 80, loss: 0.21711742877960205
step: 90, loss: 0.12673789262771606
step: 100, loss: 0.14972737431526184
step: 110, loss: 0.09037347137928009
step: 120, loss: 0.1957450956106186
step: 130, loss: 0.24330788850784302
step: 140, loss: 0.2234811633825302
step: 150, loss: 0.3233068883419037
step: 160, loss: 0.13218380510807037
step: 170, loss: 0.23453927040100098
step: 180, loss: 0.2198125123977661
step: 190, loss: 0.18057149648666382
step: 200, loss: 0.1278272271156311
step: 210, loss: 0.30770251154899597
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5475113122171946, f1=0.5246636771300449,
best_f1=0.5246636771300449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.23269188404083252
step: 10, loss: 0.29668891429901123
step: 20, loss: 0.22747352719306946
step: 30, loss: 0.10631677508354187
step: 40, loss: 0.19922329485416412
step: 50, loss: 0.17644327878952026
step: 60, loss: 0.1710863709449768
step: 70, loss: 0.08405253291130066
step: 80, loss: 0.09577640891075134
step: 90, loss: 0.15308809280395508
step: 100, loss: 0.09969162195920944
step: 110, loss: 0.12205512821674347
step: 120, loss: 0.10432372987270355
step: 130, loss: 0.2236604392528534
step: 140, loss: 0.17525604367256165
step: 150, loss: 0.22855649888515472
step: 160, loss: 0.08699771761894226
step: 170, loss: 0.12574730813503265
step: 180, loss: 0.23410123586654663
step: 190, loss: 0.2628041207790375
step: 200, loss: 0.15720342099666595
step: 210, loss: 0.2130316197872162
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5714285714285715, f1=0.5377777777777779,
best_f1=0.5377777777777779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23786941170692444
step: 10, loss: 0.11425065249204636
step: 20, loss: 0.1582907885313034
step: 30, loss: 0.0526934377849102
step: 40, loss: 0.05934862047433853
step: 50, loss: 0.0757828876376152
step: 60, loss: 0.029401257634162903
step: 70, loss: 0.33158406615257263
step: 80, loss: 0.19938619434833527
step: 90, loss: 0.019817909225821495
step: 100, loss: 0.0552256777882576
step: 110, loss: 0.10395845770835876
step: 120, loss: 0.055473167449235916
step: 130, loss: 0.1490805298089981

```

```

step: 140, loss: 0.15196362137794495
step: 150, loss: 0.1072341650724411
step: 160, loss: 0.08124972134828568
step: 170, loss: 0.051442861557006836
step: 180, loss: 0.1511572301387787
step: 190, loss: 0.14252738654613495
step: 200, loss: 0.0981842502951622
step: 210, loss: 0.030899962410330772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5777777777777777, f1=0.5407925407925408,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06418304890394211
step: 10, loss: 0.06067759171128273
step: 20, loss: 0.010540768504142761
step: 30, loss: 0.02789432555437088
step: 40, loss: 0.24668271839618683
step: 50, loss: 0.2388891875743866
step: 60, loss: 0.07766061276197433
step: 70, loss: 0.03678170591592789
step: 80, loss: 0.041031643748283386
step: 90, loss: 0.16124580800533295
step: 100, loss: 0.037645548582077026
step: 110, loss: 0.008802707307040691
step: 120, loss: 0.036165423691272736
step: 130, loss: 0.036002010107040405
step: 140, loss: 0.03776955232024193
step: 150, loss: 0.16276437044143677
step: 160, loss: 0.07168750464916229
step: 170, loss: 0.05669604241847992
step: 180, loss: 0.18283259868621826
step: 190, loss: 0.19919568300247192
step: 200, loss: 0.10446713864803314
step: 210, loss: 0.030954785645008087
/usr/local/lib/python3.7/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=0.5503875968992247, f1=0.5372549019607844,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045588418841362
step: 10, loss: 0.052130624651908875
step: 20, loss: 0.03239625319838524
step: 30, loss: 0.1892717182636261
step: 40, loss: 0.045482292771339417
step: 50, loss: 0.00640915846452117
step: 60, loss: 0.09839288145303726
step: 70, loss: 0.0036327040288597345
step: 80, loss: 0.09743139892816544
step: 90, loss: 0.1588192880153656
step: 100, loss: 0.04309890791773796
step: 110, loss: 0.02123895287513733
step: 120, loss: 0.024586156010627747
step: 130, loss: 0.02282344177365303
step: 140, loss: 0.05603974685072899
step: 150, loss: 0.028907284140586853
step: 160, loss: 0.16272515058517456
step: 170, loss: 0.014008079655468464
step: 180, loss: 0.004210271406918764
step: 190, loss: 0.09475181996822357
step: 200, loss: 0.014854051172733307
step: 210, loss: 0.004593452904373407
/usr/local/lib/python3.7/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=0.5533980582524272, f1=0.48, best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016210047528147697
step: 10, loss: 0.007145763374865055
step: 20, loss: 0.038192544132471085
step: 30, loss: 0.01864386722445488
step: 40, loss: 0.11427731812000275
step: 50, loss: 0.13584628701210022
step: 60, loss: 0.22958429157733917
step: 70, loss: 0.0159049853682518
step: 80, loss: 0.02375716157257557
step: 90, loss: 0.002501392737030983
step: 100, loss: 0.0101651381701231
step: 110, loss: 0.05009928345680237
step: 120, loss: 0.05087742581963539
step: 130, loss: 0.0072434935718774796
step: 140, loss: 0.0014046183787286282
step: 150, loss: 0.013508536852896214
step: 160, loss: 0.23155544698238373

```

```

step: 170, loss: 0.019732512533664703
step: 180, loss: 0.0049870675429701805
step: 190, loss: 0.020794441923499107
step: 200, loss: 0.010648439638316631
step: 210, loss: 0.09669598191976547
/usr/local/lib/python3.7/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=0.5468749999999999, f1=0.5304518664047151,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011078723473474383
step: 10, loss: 0.0018988187657669187
step: 20, loss: 0.034322261810302734
step: 30, loss: 0.007271449547261
step: 40, loss: 0.018008364364504814
step: 50, loss: 0.03665551543235779
step: 60, loss: 0.21861451864242554
step: 70, loss: 0.05635634809732437
step: 80, loss: 0.006119508296251297
step: 90, loss: 0.006933131255209446
step: 100, loss: 0.05215873941779137
step: 110, loss: 0.0009351029875688255
step: 120, loss: 0.008122698403894901
step: 130, loss: 0.023506898432970047
step: 140, loss: 0.0018694997997954488
step: 150, loss: 0.016262413933873177
step: 160, loss: 0.008480997756123543
step: 170, loss: 0.03710359334945679
step: 180, loss: 0.037356968969106674
step: 190, loss: 0.06958523392677307
step: 200, loss: 0.07775446027517319
step: 210, loss: 0.09244990348815918
/usr/local/lib/python3.7/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=0.5482456140350876, f1=0.5011286681715574,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.038733478635549545
step: 10, loss: 0.021527359262108803
step: 20, loss: 0.10833714157342911
step: 30, loss: 0.08457077294588089
step: 40, loss: 0.06831037998199463

```

```

step: 50, loss: 0.004076722543686628
step: 60, loss: 0.015897028148174286
step: 70, loss: 0.008440261706709862
step: 80, loss: 0.0006118903402239084
step: 90, loss: 0.030619865283370018
step: 100, loss: 0.04110957682132721
step: 110, loss: 0.00044402838102541864
step: 120, loss: 0.035575103014707565
step: 130, loss: 0.1446317434310913
step: 140, loss: 0.012801970355212688
step: 150, loss: 0.012635142542421818
step: 160, loss: 0.0019351925002411008
step: 170, loss: 0.052843689918518066
step: 180, loss: 0.025861449539661407
step: 190, loss: 0.010593120940029621
step: 200, loss: 0.04752863198518753
step: 210, loss: 0.05287764221429825
/usr/local/lib/python3.7/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=0.5450643776824035, f1=0.5287846481876332,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06557059288024902
step: 10, loss: 0.00273911957629025
step: 20, loss: 0.06050655245780945
step: 30, loss: 0.01024167612195015
step: 40, loss: 0.020097676664590836
step: 50, loss: 0.0011646089842543006
step: 60, loss: 0.022000152617692947
step: 70, loss: 0.0007652556523680687
step: 80, loss: 0.0004466715326998383
step: 90, loss: 0.013351485133171082
step: 100, loss: 0.00037710429751314223
step: 110, loss: 0.001487110392190516
step: 120, loss: 0.002808938967064023
step: 130, loss: 0.00032305510831065476
step: 140, loss: 0.000543520727660507
step: 150, loss: 0.1339767873287201
step: 160, loss: 0.019232049584388733
step: 170, loss: 0.002073881682008505
step: 180, loss: 0.0012358464300632477
step: 190, loss: 0.015193362720310688
step: 200, loss: 0.001758066238835454
step: 210, loss: 0.008526064455509186
/usr/local/lib/python3.7/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=0.5598290598290598, f1=0.5319148936170213,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016322104260325432
step: 10, loss: 0.0009581171325407922
step: 20, loss: 0.028004396706819534
step: 30, loss: 0.023574648424983025
step: 40, loss: 0.08842404186725616
step: 50, loss: 0.005129501223564148
step: 60, loss: 0.0028509011026471853
step: 70, loss: 0.0003437946434132755
step: 80, loss: 0.007254305295646191
step: 90, loss: 0.011335107497870922
step: 100, loss: 0.00338801764883101
step: 110, loss: 0.050030145794153214
step: 120, loss: 0.0035053188912570477
step: 130, loss: 0.041405171155929565
step: 140, loss: 0.006021833512932062
step: 150, loss: 0.0036095203831791878
step: 160, loss: 0.00040360819548368454
step: 170, loss: 0.1400187909603119
step: 180, loss: 0.034177228808403015
step: 190, loss: 0.009411664679646492
step: 200, loss: 0.0005271978443488479
step: 210, loss: 0.0006978139863349497
/usr/local/lib/python3.7/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=0.5454545454545455, f1=0.5253863134657837,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015427448088303208
step: 10, loss: 0.0006774357170797884
step: 20, loss: 0.001980766886845231
step: 30, loss: 0.00021782809926662594
step: 40, loss: 0.005051047541201115
step: 50, loss: 0.03279599919915199
step: 60, loss: 0.0005167412455193698
step: 70, loss: 0.0011620429577305913
step: 80, loss: 0.011038139462471008
step: 90, loss: 0.0008931384072639048
step: 100, loss: 0.004324750043451786

```

```

step: 110, loss: 0.00029363937210291624
step: 120, loss: 0.0006813525105826557
step: 130, loss: 0.0005621860036626458
step: 140, loss: 0.0021925976034253836
step: 150, loss: 0.00023741326003801078
step: 160, loss: 0.013191218487918377
step: 170, loss: 0.00032931542955338955
step: 180, loss: 0.0028330611530691385
step: 190, loss: 0.0025649252347648144
step: 200, loss: 0.07503445446491241
step: 210, loss: 0.00037512448034249246
/usr/local/lib/python3.7/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=0.5406593406593406, f1=0.5175438596491229,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013515014143195003
step: 10, loss: 0.00042542011942714453
step: 20, loss: 0.0023337367456406355
step: 30, loss: 0.004363660234957933
step: 40, loss: 0.0004033032455481589
step: 50, loss: 0.0003587840183172375
step: 60, loss: 0.0007341809687204659
step: 70, loss: 0.00044293407700024545
step: 80, loss: 0.00040514254942536354
step: 90, loss: 0.00039503644802607596
step: 100, loss: 0.01066469307988882
step: 110, loss: 0.00014311082486528903
step: 120, loss: 0.0020137629471719265
step: 130, loss: 0.0002950081252492964
step: 140, loss: 0.02674737572669983
step: 150, loss: 0.0008648482034914196
step: 160, loss: 0.00038199612754397094
step: 170, loss: 0.0002796909830067307
step: 180, loss: 0.045629508793354034
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.00037186851841397583
step: 200, loss: 0.008373904973268509
step: 210, loss: 0.0006549885147251189
/usr/local/lib/python3.7/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=0.5446623093681916, f1=0.512035010940919,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004763074219226837
step: 10, loss: 0.0001654528605286032
step: 20, loss: 0.00018799654208123684
step: 30, loss: 0.0008467945153824985
step: 40, loss: 0.0009155849111266434
step: 50, loss: 0.013336081057786942
step: 60, loss: 0.000902824045624584
step: 70, loss: 0.0036849495954811573
step: 80, loss: 0.013761667534708977
step: 90, loss: 0.0011510590557008982
step: 100, loss: 0.00041440490167587996
step: 110, loss: 0.0003608656697906554
step: 120, loss: 0.009256316348910332
step: 130, loss: 0.00014097137318458408
step: 140, loss: 0.0002466741425450891
step: 150, loss: 0.0011641581077128649
step: 160, loss: 0.0008593842503614724
step: 170, loss: 0.014475350268185139
step: 180, loss: 0.0002519182162359357
step: 190, loss: 0.0003847407642751932
step: 200, loss: 0.00042921010754071176
step: 210, loss: 0.003973004873842001
/usr/local/lib/python3.7/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=0.5417607223476298, f1=0.5080831408775981,
best_f1=0.5407925407925408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005563080194406211
step: 10, loss: 0.00499110808596015
step: 20, loss: 0.004320948850363493
step: 30, loss: 0.0005786483525298536
step: 40, loss: 0.0005303825601004064
step: 50, loss: 0.003208773210644722
step: 60, loss: 0.0013746476033702493
step: 70, loss: 0.03445575386285782
step: 80, loss: 0.0032127797603607178
step: 90, loss: 0.0001777633442543447
step: 100, loss: 0.002337130019441247
step: 110, loss: 0.00023547378077637404
step: 120, loss: 0.00073918019188568
step: 130, loss: 0.0011795539176091552
step: 140, loss: 0.0002774554886855185
step: 150, loss: 0.0015420132549479604

```



```

step: 160, loss: 0.0007910048589110374
step: 170, loss: 0.015982775017619133
step: 180, loss: 0.013398817740380764
step: 190, loss: 0.0011010699672624469
step: 200, loss: 0.045756299048662186
step: 210, loss: 0.01776997745037079
/usr/local/lib/python3.7/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=0.5412844036697246, f1=0.511520737327189,
best_f1=0.5407925407925408
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 [00:03, 619.69it/s]
load_f1 = 0.5887445887445888
real_f1 = 0.5688487584650113
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 [00:10, 408.62it/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 distilbert \
--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 distilbert \
--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: 0.8572285771369934
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.166527658700943
step: 20, loss: 0.15303045511245728
step: 30, loss: 0.5076959133148193
step: 40, loss: 0.26171597838401794
step: 50, loss: 0.31586530804634094
step: 60, loss: 0.3669007122516632
step: 70, loss: 0.17726483941078186
step: 80, loss: 0.529384195804596
step: 90, loss: 0.23107482492923737
step: 100, loss: 0.21800509095191956
step: 110, loss: 0.23538506031036377
step: 120, loss: 0.42242929339408875
step: 130, loss: 0.3387381434440613
step: 140, loss: 0.3382369577884674
step: 150, loss: 0.2767275273799896
step: 160, loss: 0.21333150565624237
step: 170, loss: 0.41346779465675354
step: 180, loss: 0.28646060824394226
step: 190, loss: 0.16869008541107178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37662337662337664, f1=0.39183673469387753,

```

```

best_f1=0.39183673469387753
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3312528431415558
step: 10, loss: 0.0414445698261261
step: 20, loss: 0.11489985883235931
step: 30, loss: 0.20345689356327057
step: 40, loss: 0.42527705430984497
step: 50, loss: 0.33428627252578735
step: 60, loss: 0.14881467819213867
step: 70, loss: 0.3083326816558838
step: 80, loss: 0.19714190065860748
step: 90, loss: 0.16278550028800964
step: 100, loss: 0.2508620619773865
step: 110, loss: 0.19034558534622192
step: 120, loss: 0.353670209646225
step: 130, loss: 0.16480329632759094
step: 140, loss: 0.2668127715587616
step: 150, loss: 0.05313541740179062
step: 160, loss: 0.17259734869003296
step: 170, loss: 0.22099781036376953
step: 180, loss: 0.19002050161361694
step: 190, loss: 0.27832409739494324
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7039106145251397, f1=0.7193460490463215,
best_f1=0.7193460490463215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1762283444404602
step: 10, loss: 0.19276881217956543
step: 20, loss: 0.027004621922969818
step: 30, loss: 0.02585270069539547
step: 40, loss: 0.054219987243413925
step: 50, loss: 0.1522800326347351
step: 60, loss: 0.1291799694299698
step: 70, loss: 0.2266535758972168
step: 80, loss: 0.15810121595859528
step: 90, loss: 0.07041753828525543
step: 100, loss: 0.20049326121807098
step: 110, loss: 0.11318523436784744

```

```

step: 120, loss: 0.03078237548470497
step: 130, loss: 0.05127941444516182
step: 140, loss: 0.1396005004644394
step: 150, loss: 0.11351985484361649
step: 160, loss: 0.056075915694236755
step: 170, loss: 0.026030918583273888
step: 180, loss: 0.06818602979183197
step: 190, loss: 0.3459535539150238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7478260869565218, f1=0.7478260869565218,
best_f1=0.7478260869565218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08398191630840302
step: 10, loss: 0.07794198393821716
step: 20, loss: 0.012767577543854713
step: 30, loss: 0.06758255511522293
step: 40, loss: 0.04249444231390953
step: 50, loss: 0.019935796037316322
step: 60, loss: 0.13226290047168732
step: 70, loss: 0.009497311897575855
step: 80, loss: 0.06627128273248672
step: 90, loss: 0.0561516135931015
step: 100, loss: 0.038970690220594406
step: 110, loss: 0.020568910986185074
step: 120, loss: 0.0872025191783905
step: 130, loss: 0.40556514263153076
step: 140, loss: 0.08228055387735367
step: 150, loss: 0.00584730040282011
step: 160, loss: 0.0839022621512413
step: 170, loss: 0.010481071658432484
step: 180, loss: 0.10552427917718887
step: 190, loss: 0.06238282099366188
/usr/local/lib/python3.7/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=0.7359550561797753, f1=0.7500000000000001,
best_f1=0.7478260869565218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.04152442514896393
step: 10, loss: 0.007395366672426462
step: 20, loss: 0.1978713572025299
step: 30, loss: 0.0313643179833889
step: 40, loss: 0.056547366082668304
step: 50, loss: 0.022074636071920395
step: 60, loss: 0.07004391402006149
step: 70, loss: 0.011876601725816727
step: 80, loss: 0.03456394374370575
step: 90, loss: 0.02461964264512062
step: 100, loss: 0.05855569243431091
step: 110, loss: 0.004143266472965479
step: 120, loss: 0.002978214295580983
step: 130, loss: 0.004539214540272951
step: 140, loss: 0.02134922333061695
step: 150, loss: 0.02081012912094593
step: 160, loss: 0.003653908148407936
step: 170, loss: 0.006908124312758446
step: 180, loss: 0.09502328932285309
step: 190, loss: 0.12348463386297226
/usr/local/lib/python3.7/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=0.735042735042735, f1=0.7293447293447294,
best_f1=0.7478260869565218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049057334661483765
step: 10, loss: 0.0043648420833051205
step: 20, loss: 0.11660487204790115
step: 30, loss: 0.07465752214193344
step: 40, loss: 0.030621204525232315
step: 50, loss: 0.012437577359378338
step: 60, loss: 0.008602609857916832
step: 70, loss: 0.010763486847281456
step: 80, loss: 0.05814998596906662
step: 90, loss: 0.006780992262065411
step: 100, loss: 0.001995754661038518
step: 110, loss: 0.08569616824388504
step: 120, loss: 0.007590505760163069
step: 130, loss: 0.0031828670762479305
step: 140, loss: 0.14350318908691406
step: 150, loss: 0.10467670112848282
step: 160, loss: 0.04395702853798866
step: 170, loss: 0.2567042112350464
step: 180, loss: 0.0074342358857393265
step: 190, loss: 0.09641558676958084

```

```

/usr/local/lib/python3.7/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=0.7423822714681441, f1=0.7582417582417582,
best_f1=0.7478260869565218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013240393251180649
step: 10, loss: 0.04477911815047264
step: 20, loss: 0.00551052438095212
step: 30, loss: 0.016112450510263443
step: 40, loss: 0.11402599513530731
step: 50, loss: 0.048712316900491714
step: 60, loss: 0.007206173613667488
step: 70, loss: 0.0008159138378687203
step: 80, loss: 0.015928564593195915
step: 90, loss: 0.0011462178081274033
step: 100, loss: 0.0022670235484838486
step: 110, loss: 0.002216336317360401
step: 120, loss: 0.056906674057245255
step: 130, loss: 0.0032935747876763344
step: 140, loss: 0.008152402937412262
step: 150, loss: 0.007644236087799072
step: 160, loss: 0.0014407576527446508
step: 170, loss: 0.01051356177777052
step: 180, loss: 0.01153141912072897
step: 190, loss: 0.04225265979766846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7588235294117647, f1=0.7826086956521738,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012545367935672402
step: 10, loss: 0.002002536319196224
step: 20, loss: 0.011173618957400322
step: 30, loss: 0.0054746391251683235
step: 40, loss: 0.02666569873690605
step: 50, loss: 0.000779474270530045
step: 60, loss: 0.001853804336860776
step: 70, loss: 0.0042276401072740555

```

```

step: 80, loss: 0.06956876814365387
step: 90, loss: 0.0017491194885224104
step: 100, loss: 0.0013848744565621018
step: 110, loss: 0.002192558953538537
step: 120, loss: 0.005855569150298834
step: 130, loss: 0.0014909267192706466
step: 140, loss: 0.002471919171512127
step: 150, loss: 0.10956414043903351
step: 160, loss: 0.0024111017119139433
step: 170, loss: 0.000755319488234818
step: 180, loss: 0.01996258832514286
step: 190, loss: 0.01244400255382061
/usr/local/lib/python3.7/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=0.6949602122015915, f1=0.7135416666666666,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035011584404855967
step: 10, loss: 0.006296938750892878
step: 20, loss: 0.010664538480341434
step: 30, loss: 0.0009455615654587746
step: 40, loss: 0.004172696266323328
step: 50, loss: 0.0029392822179943323
step: 60, loss: 0.00190789345651865
step: 70, loss: 0.0018243775703012943
step: 80, loss: 0.0006912895478308201
step: 90, loss: 0.06745772063732147
step: 100, loss: 0.0030850479379296303
step: 110, loss: 0.0034841150045394897
step: 120, loss: 0.01891104131937027
step: 130, loss: 0.001374399638734758
step: 140, loss: 0.00027817412046715617
step: 150, loss: 0.06144428998231888
step: 160, loss: 0.0019732662476599216
step: 170, loss: 0.0007838254678063095
step: 180, loss: 0.008138495497405529
step: 190, loss: 0.03738328069448471
/usr/local/lib/python3.7/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=0.7507163323782234, f1=0.7507002801120448,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.0010454733856022358
step: 10, loss: 0.015960093587636948
step: 20, loss: 0.004940188489854336
step: 30, loss: 0.000498059147503227
step: 40, loss: 0.000811676203738898
step: 50, loss: 0.0005630918894894421
step: 60, loss: 0.0033684256486594677
step: 70, loss: 0.007467522285878658
step: 80, loss: 0.003560095326974988
step: 90, loss: 0.0017851422308012843
step: 100, loss: 0.004999950993806124
step: 110, loss: 0.012896977365016937
step: 120, loss: 0.01610390655696392
step: 130, loss: 0.00032634855597279966
step: 140, loss: 0.008091447874903679
step: 150, loss: 0.00039006920997053385
step: 160, loss: 0.0004251493373885751
step: 170, loss: 0.0006824073498137295
step: 180, loss: 0.0008474920759908855
step: 190, loss: 0.0026366193778812885
/usr/local/lib/python3.7/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=0.7371273712737126, f1=0.767123287671233,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008220075978897512
step: 10, loss: 0.0011694116983562708
step: 20, loss: 0.001182946958579123
step: 30, loss: 0.004798915237188339
step: 40, loss: 0.0016605894779786468
step: 50, loss: 0.0005059342365711927
step: 60, loss: 0.12145913392305374
step: 70, loss: 0.0005271559930406511
step: 80, loss: 0.0007348673534579575
step: 90, loss: 0.003523129504173994
step: 100, loss: 0.010060416534543037
step: 110, loss: 0.0011554069351404905
step: 120, loss: 0.0012828013859689236
step: 130, loss: 0.0010305974865332246
step: 140, loss: 0.0022129444405436516
step: 150, loss: 0.0017669411608949304
step: 160, loss: 0.0007063214434310794
step: 170, loss: 0.010941781103610992
step: 180, loss: 0.003110639750957489
step: 190, loss: 0.002023986540734768

```

```

/usr/local/lib/python3.7/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=0.7566137566137565, f1=0.7574931880108992,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021563646441791207
step: 10, loss: 0.003104419680312276
step: 20, loss: 0.00031429313821718097
step: 30, loss: 0.014619192108511925
step: 40, loss: 0.001659910543821752
step: 50, loss: 0.0019547827541828156
step: 60, loss: 0.00048549118218943477
step: 70, loss: 0.0015632316935807467
step: 80, loss: 0.00041204920853488147
step: 90, loss: 0.003518176730722189
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0007632912020199001
step: 110, loss: 0.000542790163308382
step: 120, loss: 0.0005576214753091335
step: 130, loss: 0.00015044455358292907
step: 140, loss: 0.0006623062654398382
step: 150, loss: 0.00020527359447441995
step: 160, loss: 0.00852146465331316
step: 170, loss: 0.048503898084163666
step: 180, loss: 0.0008930277545005083
step: 190, loss: 0.0366722047328949
/usr/local/lib/python3.7/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=0.7554347826086956, f1=0.7458563535911602,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000579772109631449
step: 10, loss: 0.005248366855084896
step: 20, loss: 0.0013039917685091496
step: 30, loss: 0.0004087078559678048
step: 40, loss: 0.0016754032112658024
step: 50, loss: 0.0003131045668851584
step: 60, loss: 0.00034487113589420915
step: 70, loss: 0.0005834699259139597
step: 80, loss: 0.019447850063443184
step: 90, loss: 0.003195954253897071
step: 100, loss: 0.0006538864108733833

```

```

step: 110, loss: 0.0005415445775724947
step: 120, loss: 0.0005497834063135087
step: 130, loss: 0.023901794105768204
step: 140, loss: 0.00021249923156574368
step: 150, loss: 0.00018682975496631116
step: 160, loss: 0.0002515908854547888
step: 170, loss: 0.000968394975643605
step: 180, loss: 0.00042669250979088247
step: 190, loss: 0.0009093268308788538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.75, f1=0.7593582887700534, best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005498561076819897
step: 10, loss: 0.00030783412512391806
step: 20, loss: 0.0008033446501940489
step: 30, loss: 0.0007277973927557468
step: 40, loss: 0.0027012948412448168
step: 50, loss: 0.00014440716768149287
step: 60, loss: 0.0005918203969486058
step: 70, loss: 0.0004616364021785557
step: 80, loss: 0.00012587265518959612
step: 90, loss: 0.00027617427986115217
step: 100, loss: 0.00031455804128199816
step: 110, loss: 0.0015359631506726146
step: 120, loss: 0.00020072712504770607
step: 130, loss: 0.0005030371248722076
step: 140, loss: 0.00015398478717543185
step: 150, loss: 0.00021607184316962957
step: 160, loss: 0.002835685620084405
step: 170, loss: 0.0003356167580932379
step: 180, loss: 0.0001867703686002642
step: 190, loss: 0.00010107606794917956
/usr/local/lib/python3.7/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=0.7553191489361701, f1=0.7586206896551724,
best_f1=0.7826086956521738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002365314430790022
step: 10, loss: 0.0018361167749390006
step: 20, loss: 0.0002060531114693731
step: 30, loss: 0.000594475248362869

```

```

step: 40, loss: 0.0006580066401511431
step: 50, loss: 0.013061018660664558
step: 60, loss: 0.00029925958369858563
step: 70, loss: 0.0010600922396406531
step: 80, loss: 0.0006796402740292251
step: 90, loss: 0.00524901133030653
step: 100, loss: 0.00019353791140019894
step: 110, loss: 0.00035175777156837285
step: 120, loss: 0.0003427664050832391
step: 130, loss: 0.0003148522519040853
step: 140, loss: 0.002119321608915925
step: 150, loss: 0.0002954843803308904
step: 160, loss: 0.0016639436362311244
step: 170, loss: 0.00012991353287361562
step: 180, loss: 0.0003898032009601593
step: 190, loss: 0.0017920539248734713
/usr/local/lib/python3.7/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=0.7588075880758809, f1=0.7540983606557378,
best_f1=0.7826086956521738
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 [00:05, 341.57it/s]
load_f1 = 0.760806916426513
real_f1 = 0.7578347578347578
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 [00:10, 409.26it/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 distilbert \
  --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 distilbert \
  --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: 0.8545396327972412

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

step: 10, loss: 0.22907866537570953

step: 20, loss: 0.15505608916282654

step: 30, loss: 0.23894625902175903

step: 40, loss: 0.30833277106285095

step: 50, loss: 0.3711079955101013

step: 60, loss: 0.43680885434150696

step: 70, loss: 0.3100177049636841

step: 80, loss: 0.2566925585269928

step: 90, loss: 0.38708528876304626

step: 100, loss: 0.2324121594429016

step: 110, loss: 0.18636374175548553

step: 120, loss: 0.5943164825439453

step: 130, loss: 0.4103732109069824

```

step: 140, loss: 0.47683846950531006
step: 150, loss: 0.12309243530035019
step: 160, loss: 0.3445301353931427
step: 170, loss: 0.20163744688034058
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5547785547785548, f1=0.5701559020044543,
best_f1=0.5701559020044543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3374326229095459
step: 10, loss: 0.1056511253118515
step: 20, loss: 0.33770954608917236
step: 30, loss: 0.09517653286457062
step: 40, loss: 0.047636546194553375
step: 50, loss: 0.21346700191497803
step: 60, loss: 0.08700176328420639
step: 70, loss: 0.20300646126270294
step: 80, loss: 0.09234275668859482
step: 90, loss: 0.2112104445695877
step: 100, loss: 0.09839216619729996
step: 110, loss: 0.18300403654575348
step: 120, loss: 0.08249258995056152
step: 130, loss: 0.1360287368297577
step: 140, loss: 0.291046679019928
step: 150, loss: 0.13906553387641907
step: 160, loss: 0.14070278406143188
step: 170, loss: 0.06430090963840485
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7677261613691931, f1=0.7416267942583732,
best_f1=0.7416267942583732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06362592428922653
step: 10, loss: 0.07223133742809296

```

```

step: 20, loss: 0.020105626434087753
step: 30, loss: 0.1408572942018509
step: 40, loss: 0.004545548465102911
step: 50, loss: 0.0671609565615654
step: 60, loss: 0.06753759831190109
step: 70, loss: 0.04924815893173218
step: 80, loss: 0.2466868907213211
step: 90, loss: 0.12833461165428162
step: 100, loss: 0.04233267903327942
step: 110, loss: 0.15693409740924835
step: 120, loss: 0.051646772772073746
step: 130, loss: 0.16111694276332855
step: 140, loss: 0.008786926977336407
step: 150, loss: 0.0566413514316082
step: 160, loss: 0.08069819957017899
step: 170, loss: 0.08642488718032837
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7681498829039812, f1=0.7444933920704845,
best_f1=0.7444933920704845
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06277557462453842
step: 10, loss: 0.15195931494235992
step: 20, loss: 0.021009160205721855
step: 30, loss: 0.029705476015806198
step: 40, loss: 0.005788338836282492
step: 50, loss: 0.004088927526026964
step: 60, loss: 0.05667633190751076
step: 70, loss: 0.009998923167586327
step: 80, loss: 0.02409708872437477
step: 90, loss: 0.10068763047456741
step: 100, loss: 0.0119001604616642
step: 110, loss: 0.026464998722076416
step: 120, loss: 0.058984365314245224
step: 130, loss: 0.10231606662273407
step: 140, loss: 0.006486683152616024
step: 150, loss: 0.004984432365745306
step: 160, loss: 0.0425010547041893
step: 170, loss: 0.12881924211978912
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7780548628428928, f1=0.7596153846153846,
best_f1=0.7596153846153846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020636310800909996
step: 10, loss: 0.052941374480724335
step: 20, loss: 0.017785929143428802
step: 30, loss: 0.027151694521307945
step: 40, loss: 0.03864916414022446
step: 50, loss: 0.09035613387823105
step: 60, loss: 0.001988059375435114
step: 70, loss: 0.013507911935448647
step: 80, loss: 0.039125170558691025
step: 90, loss: 0.06214452162384987
step: 100, loss: 0.009743331000208855
step: 110, loss: 0.2028774470090866
step: 120, loss: 0.03393189236521721
step: 130, loss: 0.029319535940885544
step: 140, loss: 0.015118848532438278
step: 150, loss: 0.021571701392531395
step: 160, loss: 0.13694123923778534
step: 170, loss: 0.029314540326595306
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7862407862407862, f1=0.7670588235294118,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015981553122401237
step: 10, loss: 0.2196587324142456
step: 20, loss: 0.004926303401589394
step: 30, loss: 0.0195203498005867
step: 40, loss: 0.04728475213050842
step: 50, loss: 0.010826238431036472
step: 60, loss: 0.023281512781977654
step: 70, loss: 0.01540984958410263

```

```

step: 80, loss: 0.11386333405971527
step: 90, loss: 0.08650361001491547
step: 100, loss: 0.10211145132780075
step: 110, loss: 0.027511101216077805
step: 120, loss: 0.021041307598352432
step: 130, loss: 0.19796566665172577
step: 140, loss: 0.011227014474570751
step: 150, loss: 0.11928513646125793
step: 160, loss: 0.0361255519092083
step: 170, loss: 0.0038780735339969397
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.75, f1=0.755868544600939, best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012319817207753658
step: 10, loss: 0.0008352731238119304
step: 20, loss: 0.0016612838953733444
step: 30, loss: 0.003169635310769081
step: 40, loss: 0.0037059083115309477
step: 50, loss: 0.001782935461960733
step: 60, loss: 0.16599857807159424
step: 70, loss: 0.02806439995765686
step: 80, loss: 0.006299048662185669
step: 90, loss: 0.09033376723527908
step: 100, loss: 0.006874765735119581
step: 110, loss: 0.00029144450672902167
step: 120, loss: 0.17058537900447845
step: 130, loss: 0.4090683162212372
step: 140, loss: 0.026666752994060516
step: 150, loss: 0.11541066318750381
step: 160, loss: 0.006589693017303944
step: 170, loss: 0.011799967847764492
/usr/local/lib/python3.7/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=0.7523809523809523, f1=0.7563805104408352,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004534249193966389
step: 10, loss: 0.006421217694878578
step: 20, loss: 0.00713354954496026
step: 30, loss: 0.0041785514913499355
step: 40, loss: 0.0021438817493617535

```

```

step: 50, loss: 0.0003162687935400754
step: 60, loss: 0.003271977650001645
step: 70, loss: 0.012467610649764538
step: 80, loss: 0.006573855876922607
step: 90, loss: 0.10190508514642715
step: 100, loss: 0.05152124911546707
step: 110, loss: 0.0028327989857643843
step: 120, loss: 0.08677405118942261
step: 130, loss: 0.0034382983576506376
step: 140, loss: 0.008290424942970276
step: 150, loss: 0.05544883385300636
step: 160, loss: 0.02019970305263996
step: 170, loss: 0.006990419235080481
/usr/local/lib/python3.7/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=0.7409090909090909, f1=0.761904761904762,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006138266995549202
step: 10, loss: 0.0015892990631982684
step: 20, loss: 0.003177076578140259
step: 30, loss: 0.04149174317717552
step: 40, loss: 0.006928264629095793
step: 50, loss: 0.006750733591616154
step: 60, loss: 0.002856912324205041
step: 70, loss: 0.010766218416392803
step: 80, loss: 0.0421045757830143
step: 90, loss: 0.24301235377788544
step: 100, loss: 0.04567313939332962
step: 110, loss: 0.0015149860410019755
step: 120, loss: 0.013185394927859306
step: 130, loss: 0.001741337706334889
step: 140, loss: 0.0031143813394010067
step: 150, loss: 0.04271656647324562
step: 160, loss: 0.01822289265692234
step: 170, loss: 0.00906878150999546
/usr/local/lib/python3.7/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=0.7421052631578947, f1=0.7733990147783252,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004606606438755989

```

```

step: 10, loss: 0.001243271748535335
step: 20, loss: 0.0009062589379027486
step: 30, loss: 0.0003322853008285165
step: 40, loss: 0.056179311126470566
step: 50, loss: 0.00901410449296236
step: 60, loss: 0.0014260764000937343
step: 70, loss: 0.010660399682819843
step: 80, loss: 0.024226561188697815
step: 90, loss: 0.008732417598366737
step: 100, loss: 0.005644013173878193
step: 110, loss: 0.0007745674811303616
step: 120, loss: 0.021559301763772964
step: 130, loss: 0.0003308845334686339
step: 140, loss: 0.0010319782886654139
step: 150, loss: 0.00019965905812568963
step: 160, loss: 0.005193619057536125
step: 170, loss: 0.0003346198645886034
/usr/local/lib/python3.7/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=0.7360406091370559, f1=0.7707317073170732,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002506383752916008
step: 10, loss: 0.03639603778719902
step: 20, loss: 0.1653519719839096
step: 30, loss: 0.0016442316118627787
step: 40, loss: 0.003546281950548291
step: 50, loss: 0.0039533288218081
step: 60, loss: 0.0018202379578724504
step: 70, loss: 0.00114004360511899
step: 80, loss: 0.0006811315543018281
step: 90, loss: 0.0008217502036131918
step: 100, loss: 0.0014218413271009922
step: 110, loss: 0.004942712839692831
step: 120, loss: 0.17196360230445862
step: 130, loss: 0.006789408624172211
step: 140, loss: 0.01927034556865692
step: 150, loss: 0.001490470371209085
step: 160, loss: 0.0012440626742318273
step: 170, loss: 0.006149684078991413
/usr/local/lib/python3.7/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=0.7355163727959698, f1=0.7699757869249395,
best_f1=0.7670588235294118

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006547577679157257
step: 10, loss: 0.0013359565054997802
step: 20, loss: 0.023720787838101387
step: 30, loss: 0.0013691646745428443
step: 40, loss: 0.004953958094120026
step: 50, loss: 0.010712460614740849
step: 60, loss: 0.08988359570503235
step: 70, loss: 0.10248813033103943
step: 80, loss: 0.07066585123538971
step: 90, loss: 0.0027373419143259525
step: 100, loss: 0.001206892542541027
step: 110, loss: 0.00019097790936939418
step: 120, loss: 0.006647647824138403
step: 130, loss: 0.0010154402116313577
step: 140, loss: 0.001848651678301394
step: 150, loss: 0.00022104136587586254
step: 160, loss: 0.03508903086185455
step: 170, loss: 0.004536136984825134
/usr/local/lib/python3.7/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=0.7244897959183674, f1=0.788177339901478,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017944228602573276
step: 10, loss: 0.038003336638212204
step: 20, loss: 0.0006388889160007238
step: 30, loss: 0.00016031417180784047
step: 40, loss: 0.000166332843946293
step: 50, loss: 0.00017375347670167685
step: 60, loss: 0.00019356115080881864
step: 70, loss: 0.0025137572083622217
step: 80, loss: 0.00032474129693582654
step: 90, loss: 0.00016424732166342437
step: 100, loss: 0.00043955029104836285
step: 110, loss: 0.0003407914482522756
step: 120, loss: 0.00015086954226717353
step: 130, loss: 0.00013729653437621891
step: 140, loss: 0.00010715433745644987
step: 150, loss: 0.006848642602562904
step: 160, loss: 0.0003000669239554554
step: 170, loss: 0.0023153782822191715
/usr/local/lib/python3.7/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=0.7416267942583732, f1=0.7603686635944701,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009848017944023013
step: 10, loss: 0.0004520628717727959
step: 20, loss: 0.0007262186263687909
step: 30, loss: 0.23276615142822266
step: 40, loss: 0.062498338520526886
step: 50, loss: 0.20433346927165985
step: 60, loss: 0.0001626661396585405
step: 70, loss: 0.0006445432081818581
step: 80, loss: 0.1300298571586609
step: 90, loss: 0.0004091049777343869
step: 100, loss: 0.001449784031137824
step: 110, loss: 0.002366352127864957
step: 120, loss: 0.01720239222049713
step: 130, loss: 0.0014475618954747915
step: 140, loss: 0.0006271878373809159
step: 150, loss: 0.0005346942343749106
step: 160, loss: 0.02014835551381111
step: 170, loss: 0.003081131260842085
/usr/local/lib/python3.7/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=0.7543424317617866, f1=0.7932692307692307,
best_f1=0.7670588235294118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007874438888393342
step: 10, loss: 0.0007922201766632497
step: 20, loss: 0.000798744207713753
step: 30, loss: 0.03133125603199005
step: 40, loss: 0.0002963104343507439
step: 50, loss: 0.00212344853207469
step: 60, loss: 0.0008606077753938735
step: 70, loss: 0.030612042173743248
step: 80, loss: 0.0009749546297825873
step: 90, loss: 0.00014048724551685154
step: 100, loss: 0.0011232170509174466
step: 110, loss: 0.00044428594992496073
step: 120, loss: 0.024623753502964973
step: 130, loss: 0.0005238060257397592
step: 140, loss: 0.0003301195683889091

```

```

step: 150, loss: 0.019399745389819145
step: 160, loss: 0.00036027427995577455
step: 170, loss: 0.00019023180357180536
/usr/local/lib/python3.7/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=0.743142144638404, f1=0.7961165048543688,
best_f1=0.7670588235294118
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 [00:04, 433.19it/s]
load_f1 = 0.7808564231738037
real_f1 = 0.7791563275434243
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 [00:12, 338.36it/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 distilbert \
--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 distilbert \
--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: 0.8046383857727051
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4903336465358734
step: 20, loss: 0.5924717783927917
step: 30, loss: 0.4965613782405853
step: 40, loss: 0.37110212445259094
step: 50, loss: 0.2515426576137543
step: 60, loss: 0.24534407258033752
step: 70, loss: 0.11003642529249191
step: 80, loss: 0.2685457170009613
step: 90, loss: 0.07567218691110611
step: 100, loss: 0.1405859887599945
step: 110, loss: 0.08670969307422638
step: 120, loss: 0.021721096709370613
step: 130, loss: 0.0257062129676342
step: 140, loss: 0.14188988506793976
step: 150, loss: 0.16156788170337677
step: 160, loss: 0.24404819309711456
step: 170, loss: 0.024869687855243683
step: 180, loss: 0.0059579890221357346
step: 190, loss: 0.044815484434366226
step: 200, loss: 0.04116314649581909
step: 210, loss: 0.01881684921681881
step: 220, loss: 0.00855773501098156
step: 230, loss: 0.05538474768400192
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9521739130434783, f1=0.9477124183006536,
best_f1=0.9477124183006536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03107597678899765
step: 10, loss: 0.04572349041700363
step: 20, loss: 0.034351371228694916
step: 30, loss: 0.01204507052898407
step: 40, loss: 0.027147410437464714
step: 50, loss: 0.03646870329976082
step: 60, loss: 0.060280222445726395
step: 70, loss: 0.04239841178059578
step: 80, loss: 0.011930586770176888
step: 90, loss: 0.06013954430818558
step: 100, loss: 0.21064575016498566
step: 110, loss: 0.06932034343481064
step: 120, loss: 0.07974683493375778
step: 130, loss: 0.02336261235177517
step: 140, loss: 0.04130816459655762
step: 150, loss: 0.03242555260658264
step: 160, loss: 0.014561380259692669
step: 170, loss: 0.06431681662797928
step: 180, loss: 0.04977152496576309
step: 190, loss: 0.14121420681476593
step: 200, loss: 0.04900098219513893
step: 210, loss: 0.03847672417759895
step: 220, loss: 0.003719762898981571
step: 230, loss: 0.03706414997577667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9564270152505447, f1=0.9478260869565216,
best_f1=0.9478260869565216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0909862071275711
step: 10, loss: 0.011484317481517792

```

```

step: 20, loss: 0.01083631906658411
step: 30, loss: 0.01613752171397209
step: 40, loss: 0.13883911073207855
step: 50, loss: 0.003543444909155369
step: 60, loss: 0.050060342997312546
step: 70, loss: 0.005736449267715216
step: 80, loss: 0.026395289227366447
step: 90, loss: 0.16933126747608185
step: 100, loss: 0.004140493925660849
step: 110, loss: 0.0688500702381134
step: 120, loss: 0.013685124926269054
step: 130, loss: 0.005175324156880379
step: 140, loss: 0.005598198156803846
step: 150, loss: 0.004397256765514612
step: 160, loss: 0.013243754394352436
step: 170, loss: 0.004212826956063509
step: 180, loss: 0.02135605551302433
step: 190, loss: 0.00517387967556715
step: 200, loss: 0.05193119868636131
step: 210, loss: 0.021894140169024467
step: 220, loss: 0.005891463253647089
step: 230, loss: 0.03718996047973633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9623059866962307, f1=0.9556541019955654,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012206224724650383
step: 10, loss: 0.0031858014408499002
step: 20, loss: 0.0008050522883422673
step: 30, loss: 0.000996192218735814
step: 40, loss: 0.007219039835035801
step: 50, loss: 0.0012111482210457325
step: 60, loss: 0.0008632897515781224
step: 70, loss: 0.01991378515958786
step: 80, loss: 0.11396275460720062
step: 90, loss: 0.0764181837439537
step: 100, loss: 0.005270565859973431
step: 110, loss: 0.030874798074364662
step: 120, loss: 0.021109892055392265
step: 130, loss: 0.0020001560915261507

```

```

step: 140, loss: 0.004582713823765516
step: 150, loss: 0.002925907028838992
step: 160, loss: 0.017156118527054787
step: 170, loss: 0.01740744523704052
step: 180, loss: 0.04742579907178879
step: 190, loss: 0.07301410287618637
step: 200, loss: 0.002534370869398117
step: 210, loss: 0.003745294874534011
step: 220, loss: 0.021078504621982574
step: 230, loss: 0.0003958141023758799
/usr/local/lib/python3.7/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=0.9590254706533776, f1=0.9534368070953436,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026277629658579826
step: 10, loss: 0.00643490394577384
step: 20, loss: 0.00025879754684865475
step: 30, loss: 0.0025639608502388
step: 40, loss: 0.0009927237406373024
step: 50, loss: 0.00023881309607531875
step: 60, loss: 0.0004181672993581742
step: 70, loss: 0.00012141672777943313
step: 80, loss: 0.0004271502257324755
step: 90, loss: 0.0008937979000620544
step: 100, loss: 0.11024025082588196
step: 110, loss: 0.0006072326796129346
step: 120, loss: 0.030635124072432518
step: 130, loss: 0.09509008377790451
step: 140, loss: 0.029309822246432304
step: 150, loss: 0.00023563954164274037
step: 160, loss: 0.11550644040107727
step: 170, loss: 0.014034162275493145
step: 180, loss: 0.0006349787581712008
step: 190, loss: 0.003355169901624322
step: 200, loss: 0.0007463152287527919
step: 210, loss: 0.0003581151831895113
step: 220, loss: 0.003329802770167589
step: 230, loss: 0.010629064403474331
/usr/local/lib/python3.7/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=0.9587513935340021, f1=0.9597315436241611,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06690080463886261
step: 10, loss: 0.023378606885671616
step: 20, loss: 0.004412560258060694
step: 30, loss: 0.0008206978673115373
step: 40, loss: 0.0046699922531843185
step: 50, loss: 0.013107651844620705
step: 60, loss: 0.003256098134443164
step: 70, loss: 0.00195965263992548
step: 80, loss: 0.0007261214195750654
step: 90, loss: 0.025647563859820366
step: 100, loss: 0.0015900885919108987
step: 110, loss: 0.05871474742889404
step: 120, loss: 0.00046217686031013727
step: 130, loss: 0.00020923223928548396
step: 140, loss: 0.0008432668983004987
step: 150, loss: 0.0016416388098150492
step: 160, loss: 0.016191573813557625
step: 170, loss: 0.0006789572653360665
step: 180, loss: 0.0006034052348695695
step: 190, loss: 0.0006072503747418523
step: 200, loss: 0.017904436215758324
step: 210, loss: 0.0005143631133250892
step: 220, loss: 0.0011470186291262507
step: 230, loss: 0.0002676590229384601
/usr/local/lib/python3.7/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=0.9583333333333334, f1=0.9582417582417582,
best_f1=0.9556541019955654
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13577322661876678
step: 10, loss: 0.0021649200934916735
step: 20, loss: 0.0004636141820810735
step: 30, loss: 0.0019853597041219473
step: 40, loss: 0.0005687936209142208
step: 50, loss: 0.0003677413915283978
step: 60, loss: 0.008668274618685246
step: 70, loss: 0.0006159887416288257
step: 80, loss: 0.003914670553058386
step: 90, loss: 0.007527743000537157
step: 100, loss: 0.0006953478441573679
step: 110, loss: 0.00023318451712839305
step: 120, loss: 0.0011783005902543664
step: 130, loss: 0.000962116930168122

```

```

step: 140, loss: 0.00011680062743835151
step: 150, loss: 0.0004936306504532695
step: 160, loss: 0.0009092445252463222
step: 170, loss: 0.00019137165509164333
step: 180, loss: 0.000887383590452373
step: 190, loss: 0.00011671109677990898
step: 200, loss: 0.0015521316090598702
step: 210, loss: 0.0002200257295044139
step: 220, loss: 0.00010714754898799583
step: 230, loss: 0.003794429823756218
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9632107023411371, f1=0.9621380846325166,
best_f1=0.9621380846325166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009313217364251614
step: 10, loss: 0.05142316222190857
step: 20, loss: 0.0015151083935052156
step: 30, loss: 0.0024704912211745977
step: 40, loss: 0.00015085938503034413
step: 50, loss: 0.0011300835758447647
step: 60, loss: 0.0001947565033333376
step: 70, loss: 0.00010870342521229759
step: 80, loss: 0.0001504178944742307
step: 90, loss: 9.586701344233006e-05
step: 100, loss: 0.00013107513950672
step: 110, loss: 0.0003094378625974059
step: 120, loss: 0.0015509964432567358
step: 130, loss: 0.0004903741646558046
step: 140, loss: 6.712411413900554e-05
step: 150, loss: 0.0028937675524502993
step: 160, loss: 0.00019097117183264345
step: 170, loss: 9.197311010211706e-05
step: 180, loss: 0.000686357612721622
step: 190, loss: 0.0011933420319110155
step: 200, loss: 0.026459351181983948
step: 210, loss: 8.459360105916858e-05
step: 220, loss: 0.00012786882871296257
step: 230, loss: 0.0538397915661335
/usr/local/lib/python3.7/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=0.961625282167043, f1=0.9569160997732427,
best_f1=0.9621380846325166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001291436783503741
step: 10, loss: 0.0015528511721640825
step: 20, loss: 0.004799242597073317
step: 30, loss: 0.0019257600652053952
step: 40, loss: 5.8650595747167245e-05
step: 50, loss: 0.0006128186942078173
step: 60, loss: 0.00038767687510699034
step: 70, loss: 0.0002676775911822915
step: 80, loss: 0.025077326223254204
step: 90, loss: 0.005957814399152994
step: 100, loss: 0.01064020674675703
step: 110, loss: 0.00016186815628316253
step: 120, loss: 0.028889812529087067
step: 130, loss: 8.802070806268603e-05
step: 140, loss: 0.029735371470451355
step: 150, loss: 0.00039576657582074404
step: 160, loss: 0.0003336210211273283
step: 170, loss: 5.8325309510109946e-05
step: 180, loss: 0.0011302991770207882
step: 190, loss: 0.00042556834523566067
step: 200, loss: 0.0014470387250185013
step: 210, loss: 0.0009179693879559636
step: 220, loss: 0.03150651976466179
step: 230, loss: 0.00012666668044403195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9658213891951488, f1=0.9591160220994475,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012017534754704684
step: 10, loss: 8.783108205534518e-05
step: 20, loss: 0.0001358946756226942
step: 30, loss: 0.0001243295264430344
step: 40, loss: 6.685112020932138e-05
step: 50, loss: 0.004068133421242237

```

```

step: 60, loss: 0.012710759416222572
step: 70, loss: 8.813187014311552e-05
step: 80, loss: 0.00036931733484379947
step: 90, loss: 6.730736640747637e-05
step: 100, loss: 0.001936982269398868
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 110, loss: 0.06132485345005989
step: 120, loss: 0.018570616841316223
step: 130, loss: 0.0001687201001914218
step: 140, loss: 0.007226430345326662
step: 150, loss: 0.0009918054565787315
step: 160, loss: 7.499866478610784e-05
step: 170, loss: 0.008666926063597202
step: 180, loss: 0.0004227555182296783
step: 190, loss: 9.956947906175628e-05
step: 200, loss: 5.1909391913795844e-05
step: 210, loss: 6.939913146197796e-05
step: 220, loss: 0.0001892577129183337
step: 230, loss: 0.0002691195113584399
/usr/local/lib/python3.7/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=0.961883408071749, f1=0.9605411499436302,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024592794943600893
step: 10, loss: 0.00025223413831554353
step: 20, loss: 7.74441214161925e-05
step: 30, loss: 0.0002581859880592674
step: 40, loss: 0.004758692812174559
step: 50, loss: 0.001324064563959837
step: 60, loss: 0.0001329543738393113
step: 70, loss: 0.0002826550044119358
step: 80, loss: 0.0003562209603842348
step: 90, loss: 0.0003092079423367977
step: 100, loss: 0.0002957773394882679
step: 110, loss: 0.00014971772907301784
step: 120, loss: 0.0003211955481674522
step: 130, loss: 0.005562921985983849
step: 140, loss: 8.252388215623796e-05
step: 150, loss: 8.863874245435e-05
step: 160, loss: 0.0013480441411957145
step: 170, loss: 0.12829110026359558
step: 180, loss: 0.003557612420991063
step: 190, loss: 0.00041937772766686976
step: 200, loss: 0.0002604954061098397

```



```

step: 210, loss: 5.951529601588845e-05
step: 220, loss: 0.00021723810641560704
step: 230, loss: 0.014053362421691418
/usr/local/lib/python3.7/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=0.9645232815964524, f1=0.9566184649610678,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004667996254283935
step: 10, loss: 0.00028314816881902516
step: 20, loss: 0.0021381762344390154
step: 30, loss: 0.026275834068655968
step: 40, loss: 4.642194107873365e-05
step: 50, loss: 7.85894226282835e-05
step: 60, loss: 0.017862217500805855
step: 70, loss: 5.6037777540041134e-05
step: 80, loss: 0.00010220378317171708
step: 90, loss: 8.28912016004324e-05
step: 100, loss: 7.841311889933422e-05
step: 110, loss: 8.117916877381504e-05
step: 120, loss: 6.390204362105578e-05
step: 130, loss: 0.00014784974337089807
step: 140, loss: 9.091177344089374e-05
step: 150, loss: 3.2472478778799996e-05
step: 160, loss: 0.02019141986966133
step: 170, loss: 7.54656721255742e-05
step: 180, loss: 6.684243999188766e-05
step: 190, loss: 0.00012751694885082543
step: 200, loss: 0.00020827124535571784
step: 210, loss: 8.065587462624535e-05
step: 220, loss: 0.02975805290043354
step: 230, loss: 0.00011280538456048816
/usr/local/lib/python3.7/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=0.9655172413793103, f1=0.9553571428571428,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.946795423980802e-05
step: 10, loss: 5.301769488141872e-05
step: 20, loss: 0.00024917040718719363
step: 30, loss: 0.019834451377391815
step: 40, loss: 0.00033000411349348724

```

```

step: 50, loss: 0.00012807243911083788
step: 60, loss: 0.00011844051914522424
step: 70, loss: 5.673519626725465e-05
step: 80, loss: 6.347990711219609e-05
step: 90, loss: 3.3764885301934555e-05
step: 100, loss: 0.00010103835666086525
step: 110, loss: 0.00017668014334049076
step: 120, loss: 5.388250792748295e-05
step: 130, loss: 0.00212054792791605
step: 140, loss: 0.0134204076603055
step: 150, loss: 8.431362948613241e-05
step: 160, loss: 4.761311356560327e-05
step: 170, loss: 7.23331977496855e-05
step: 180, loss: 0.00013031158596277237
step: 190, loss: 6.787176243960857e-05
step: 200, loss: 5.466748552862555e-05
step: 210, loss: 9.313788177678362e-05
step: 220, loss: 4.937150879413821e-05
step: 230, loss: 5.3805943025508896e-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=0.9678135405105438, f1=0.9577777777777777,
best_f1=0.9577777777777777
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.993705831817351e-05
step: 10, loss: 3.864525569952093e-05
step: 20, loss: 0.0013838771265000105
step: 30, loss: 6.947456859052181e-05
step: 40, loss: 4.191720290691592e-05
step: 50, loss: 7.500416541006416e-05
step: 60, loss: 8.765904931351542e-05
step: 70, loss: 8.572621300118044e-05
step: 80, loss: 8.153299131663516e-05
step: 90, loss: 7.520365033997223e-05
step: 100, loss: 0.0012804210418835282
step: 110, loss: 0.0005726718227379024
step: 120, loss: 0.0009379174443893135
step: 130, loss: 6.883191963424906e-05
step: 140, loss: 3.9337199268629774e-05
step: 150, loss: 5.1167931815143675e-05
step: 160, loss: 3.22823143505957e-05

```

```

step: 170, loss: 4.0163293306250125e-05
step: 180, loss: 9.967199002858251e-05
step: 190, loss: 0.0012120752362534404
step: 200, loss: 7.267783803399652e-05
step: 210, loss: 0.0005963295116089284
step: 220, loss: 0.00031355355167761445
step: 230, loss: 6.041867527528666e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9686800894854586, f1=0.9619686800894856,
best_f1=0.9619686800894856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.1095328924711794e-05
step: 10, loss: 0.0002022337430389598
step: 20, loss: 7.447702955687419e-05
step: 30, loss: 6.665789260296151e-05
step: 40, loss: 0.00011080252443207428
step: 50, loss: 6.21034560026601e-05
step: 60, loss: 3.7511923437705263e-05
step: 70, loss: 9.544532076688483e-05
step: 80, loss: 0.00011875028576469049
step: 90, loss: 3.689270306495018e-05
step: 100, loss: 4.982656901120208e-05
step: 110, loss: 8.181917655747384e-05
step: 120, loss: 0.00022068808902986348
step: 130, loss: 7.151088357204571e-05
step: 140, loss: 0.025014305487275124
step: 150, loss: 6.494394619949162e-05
step: 160, loss: 0.00016382108151447028
step: 170, loss: 5.3197905799606815e-05
step: 180, loss: 5.135338506079279e-05
step: 190, loss: 6.691722228424624e-05
step: 200, loss: 3.559029210009612e-05
step: 210, loss: 0.0018670324934646487
step: 220, loss: 0.0029686593916267157
step: 230, loss: 6.426434993045405e-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 15: dev_f1=0.9698324022346367, f1=0.9598214285714285,
best_f1=0.9598214285714285
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:08, 297.91it/s]
load_f1 = 0.9686800894854586
real_f1 = 0.9675977653631285
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 [00:12, 360.33it/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 distilbert \
--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 distilbert \
--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: 0.7943519353866577
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4523506164550781
step: 20, loss: 0.49603721499443054
step: 30, loss: 0.4817029535770416
step: 40, loss: 0.4737206995487213
step: 50, loss: 0.39937713742256165
step: 60, loss: 0.5063918232917786
step: 70, loss: 0.1397455930709839
step: 80, loss: 0.31220343708992004
step: 90, loss: 0.37663742899894714
step: 100, loss: 0.2506047189235687
step: 110, loss: 0.09224937111139297
step: 120, loss: 0.1213613748550415
step: 130, loss: 0.06089707836508751
step: 140, loss: 0.27322065830230713
step: 150, loss: 0.0616125762462616
step: 160, loss: 0.15490292012691498
step: 170, loss: 0.3204871118068695
step: 180, loss: 0.14193607866764069
step: 190, loss: 0.09087347984313965
step: 200, loss: 0.15488684177398682
step: 210, loss: 0.09613353759050369
step: 220, loss: 0.11752115935087204
step: 230, loss: 0.17440617084503174
step: 240, loss: 0.09481576830148697
step: 250, loss: 0.08159855753183365
step: 260, loss: 0.040164534002542496
step: 270, loss: 0.12744808197021484

```

```

step: 280, loss: 0.18674106895923615
step: 290, loss: 0.05582589656114578
step: 300, loss: 0.2964995801448822
step: 310, loss: 0.12154386937618256
step: 320, loss: 0.07132405787706375
step: 330, loss: 0.2083374559879303
step: 340, loss: 0.2364344298839569
step: 350, loss: 0.077774278819561
step: 360, loss: 0.08079886436462402
step: 370, loss: 0.09057862311601639
step: 380, loss: 0.29185160994529724
step: 390, loss: 0.029750710353255272
step: 400, loss: 0.05546782910823822
step: 410, loss: 0.09557216614484787
step: 420, loss: 0.10787563025951385
step: 430, loss: 0.09051328897476196
step: 440, loss: 0.143899604678154
step: 450, loss: 0.03931334987282753
step: 460, loss: 0.025658851489424706
step: 470, loss: 0.3795336186885834
step: 480, loss: 0.15096893906593323
step: 490, loss: 0.11509127914905548
step: 500, loss: 0.09383150190114975
step: 510, loss: 0.15314997732639313
step: 520, loss: 0.16962392628192902
step: 530, loss: 0.03742473945021629
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9047169811320754, f1=0.9092618711800659,
best_f1=0.9092618711800659
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06766179203987122
step: 10, loss: 0.1878446787595749
step: 20, loss: 0.06235795095562935
step: 30, loss: 0.07640049606561661
step: 40, loss: 0.00626192893832922
step: 50, loss: 0.04787726700305939
step: 60, loss: 0.07942177355289459
step: 70, loss: 0.22996264696121216
step: 80, loss: 0.03010454773902893
step: 90, loss: 0.045484352856874466

```

```
step: 100, loss: 0.22329604625701904
step: 110, loss: 0.05631764978170395
step: 120, loss: 0.050518203526735306
step: 130, loss: 0.02941708266735077
step: 140, loss: 0.024234825745224953
step: 150, loss: 0.06896104663610458
step: 160, loss: 0.06513561308383942
step: 170, loss: 0.23864267766475677
step: 180, loss: 0.03047901764512062
step: 190, loss: 0.03929920122027397
step: 200, loss: 0.02281496301293373
step: 210, loss: 0.03461086377501488
step: 220, loss: 0.18085061013698578
step: 230, loss: 0.019019603729248047
step: 240, loss: 0.15734128654003143
step: 250, loss: 0.16324765980243683
step: 260, loss: 0.04422708973288536
step: 270, loss: 0.07708849757909775
step: 280, loss: 0.41777509450912476
step: 290, loss: 0.20077002048492432
step: 300, loss: 0.05276171490550041
step: 310, loss: 0.09203969687223434
step: 320, loss: 0.1508033722639084
step: 330, loss: 0.023333190008997917
step: 340, loss: 0.009278272278606892
step: 350, loss: 0.02547428384423256
step: 360, loss: 0.0776071697473526
step: 370, loss: 0.03291481360793114
step: 380, loss: 0.09540940076112747
step: 390, loss: 0.024887533858418465
step: 400, loss: 0.12946374714374542
step: 410, loss: 0.017755666747689247
step: 420, loss: 0.05285046622157097
step: 430, loss: 0.017115116119384766
step: 440, loss: 0.00782057736068964
step: 450, loss: 0.050054147839546204
step: 460, loss: 0.305059015750885
step: 470, loss: 0.11017963290214539
step: 480, loss: 0.24372610449790955
step: 490, loss: 0.06498382240533829
step: 500, loss: 0.023157941177487373
step: 510, loss: 0.06379688531160355
step: 520, loss: 0.004998455755412579
step: 530, loss: 0.09316748380661011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```


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=0.9155389640690621, f1=0.916319926028664,
best_f1=0.916319926028664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

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

```
step: 0, loss: 0.06807541102170944
step: 10, loss: 0.04747559502720833
step: 20, loss: 0.19379907846450806
step: 30, loss: 0.11327394843101501
step: 40, loss: 0.008811562322080135
step: 50, loss: 0.01735837012529373
step: 60, loss: 0.0048561678268015385
step: 70, loss: 0.018689338117837906
step: 80, loss: 0.003650825237855315
step: 90, loss: 0.02614516392350197
step: 100, loss: 0.023033451288938522
step: 110, loss: 0.02698150835931301
step: 120, loss: 0.06710793077945709
step: 130, loss: 0.04007217660546303
step: 140, loss: 0.03375866636633873
step: 150, loss: 0.017605677247047424
step: 160, loss: 0.005559810437262058
step: 170, loss: 0.013678023591637611
step: 180, loss: 0.013535172678530216
step: 190, loss: 0.007129790261387825
step: 200, loss: 0.03145357593894005
step: 210, loss: 0.13233228027820587
step: 220, loss: 0.06898902356624603
step: 230, loss: 0.0882113054394722
step: 240, loss: 0.022359224036335945
step: 250, loss: 0.012357868254184723
step: 260, loss: 0.002626839792355895
step: 270, loss: 0.0038875655736774206
step: 280, loss: 0.07102461904287338
step: 290, loss: 0.0535852313041687
step: 300, loss: 0.11209136247634888
step: 310, loss: 0.208004429936409
step: 320, loss: 0.1872147023677826
step: 330, loss: 0.0020913793705403805
step: 340, loss: 0.02431541495025158
step: 350, loss: 0.13518346846103668
step: 360, loss: 0.01625720039010048
step: 370, loss: 0.05483105033636093
step: 380, loss: 0.012653935700654984
step: 390, loss: 0.02731133997440338
```

```

step: 400, loss: 0.009733538143336773
step: 410, loss: 0.03314296156167984
step: 420, loss: 0.024623380973935127
step: 430, loss: 0.010936550796031952
step: 440, loss: 0.10004115104675293
step: 450, loss: 0.06200201064348221
step: 460, loss: 0.07692212611436844
step: 470, loss: 0.02193276211619377
step: 480, loss: 0.0031932832207530737
step: 490, loss: 0.14168410003185272
step: 500, loss: 0.0571889691054821
step: 510, loss: 0.008096757344901562
step: 520, loss: 0.007025676313787699
step: 530, loss: 0.0535486601293087
/usr/local/lib/python3.7/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=0.9144676979071884, f1=0.9100817438692098,
best_f1=0.916319926028664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00904729962348938
step: 10, loss: 0.011141528375446796
step: 20, loss: 0.027918698266148567
step: 30, loss: 0.00852117408066988
step: 40, loss: 0.00257376697845757
step: 50, loss: 0.013684101402759552
step: 60, loss: 0.0015169854741543531
step: 70, loss: 0.1868748813867569
step: 80, loss: 0.04433899000287056
step: 90, loss: 0.019538646563887596
step: 100, loss: 0.013552702032029629
step: 110, loss: 0.00592777831479907
step: 120, loss: 0.01391805149614811
step: 130, loss: 0.0032948763109743595
step: 140, loss: 0.006112614646553993
step: 150, loss: 0.0005956271197646856
step: 160, loss: 0.0020563837606459856
step: 170, loss: 0.011455644853413105
step: 180, loss: 0.013583547435700893
step: 190, loss: 0.01976754516363144
step: 200, loss: 0.0022561189252883196
step: 210, loss: 0.011985642835497856
step: 220, loss: 0.005236433818936348
step: 230, loss: 0.2568543553352356
step: 240, loss: 0.003517396515235305
step: 250, loss: 0.030024120584130287

```

```

step: 260, loss: 0.011892099864780903
step: 270, loss: 0.00944282952696085
step: 280, loss: 0.016730230301618576
step: 290, loss: 0.10573042929172516
step: 300, loss: 0.0013937499606981874
step: 310, loss: 0.0029797423630952835
step: 320, loss: 0.024605926126241684
step: 330, loss: 0.007397571112960577
step: 340, loss: 0.0021484948229044676
step: 350, loss: 0.005375054199248552
step: 360, loss: 0.029465602710843086
step: 370, loss: 0.023153848946094513
step: 380, loss: 0.005204758141189814
step: 390, loss: 0.1342737078666687
step: 400, loss: 0.102778859436512
step: 410, loss: 0.024018719792366028
step: 420, loss: 0.03545381501317024
step: 430, loss: 0.02334207110106945
step: 440, loss: 0.06882733851671219
step: 450, loss: 0.03612356632947922
step: 460, loss: 0.007456745952367783
step: 470, loss: 0.004991882015019655
step: 480, loss: 0.021025525406003
step: 490, loss: 0.011862984858453274
step: 500, loss: 0.027030285447835922
step: 510, loss: 0.020225299522280693
step: 520, loss: 0.11879158020019531
step: 530, loss: 0.005227312445640564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.916970802919708, f1=0.9138712601994561,
best_f1=0.9138712601994561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0046510216780006886
step: 10, loss: 0.05278909578919411
step: 20, loss: 0.018184373155236244
step: 30, loss: 0.0012204190716147423
step: 40, loss: 0.018442878499627113
step: 50, loss: 0.009962551295757294
step: 60, loss: 0.04933653026819229
step: 70, loss: 0.001918732188642025

```

step: 80, loss: 0.0035530030727386475
step: 90, loss: 0.0056816525757312775
step: 100, loss: 0.01045847125351429
step: 110, loss: 0.0009842379949986935
step: 120, loss: 0.005077371373772621
step: 130, loss: 0.002477984642609954
step: 140, loss: 0.007622147910296917
step: 150, loss: 0.0011076705995947123
step: 160, loss: 0.018344439566135406
step: 170, loss: 0.029229944571852684
step: 180, loss: 0.021022768691182137
step: 190, loss: 0.002412231173366308
step: 200, loss: 0.0008460658136755228
step: 210, loss: 0.0034969751723110676
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.0021735678892582655
step: 230, loss: 0.003338361158967018
step: 240, loss: 0.020244082435965538
step: 250, loss: 0.00397466542199254
step: 260, loss: 0.015114301815629005
step: 270, loss: 0.06949760019779205
step: 280, loss: 0.013490252196788788
step: 290, loss: 0.001538111362606287
step: 300, loss: 0.003129385644569993
step: 310, loss: 0.029682600870728493
step: 320, loss: 0.01483994536101818
step: 330, loss: 0.0005751998396590352
step: 340, loss: 0.0012234895257279277
step: 350, loss: 0.01493875589221716
step: 360, loss: 0.026437057182192802
step: 370, loss: 0.0033196539152413607
step: 380, loss: 0.004876383114606142
step: 390, loss: 0.0008417901699431241
step: 400, loss: 0.03176717087626457
step: 410, loss: 0.0010104234097525477
step: 420, loss: 0.0006807273020967841
step: 430, loss: 0.0037961460184305906
step: 440, loss: 0.015195987187325954
step: 450, loss: 0.02163030207157135
step: 460, loss: 0.07722500711679459
step: 470, loss: 0.0010393948759883642
step: 480, loss: 0.01170874759554863
step: 490, loss: 0.01948240026831627
step: 500, loss: 0.022459007799625397
step: 510, loss: 0.0036817146465182304
step: 520, loss: 0.0062103671953082085
step: 530, loss: 0.0017414214089512825
/usr/local/lib/python3.7/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=0.916206261510129, f1=0.9101741521539871,
best_f1=0.9138712601994561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006561414920724928
step: 10, loss: 0.006015758495777845
step: 20, loss: 0.0003115088620688766
step: 30, loss: 0.06769656389951706
step: 40, loss: 0.006518933456391096
step: 50, loss: 0.0005229414673522115
step: 60, loss: 0.0006120551261119545
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.013090400956571102
step: 80, loss: 0.0008978573605418205
step: 90, loss: 0.0007851833361200988
step: 100, loss: 0.000820089306216687
step: 110, loss: 0.002608120907098055
step: 120, loss: 0.026691557839512825
step: 130, loss: 0.004966315813362598
step: 140, loss: 0.0021926139015704393
step: 150, loss: 0.030105430632829666
step: 160, loss: 0.0016476755263283849
step: 170, loss: 0.00021987636864650995
step: 180, loss: 0.00047374016139656305
step: 190, loss: 0.06543130427598953
step: 200, loss: 0.004184462130069733
step: 210, loss: 0.0011294173309579492
step: 220, loss: 0.0005617480492219329
step: 230, loss: 0.0023541434202343225
step: 240, loss: 0.0007546855485998094
step: 250, loss: 0.0005344762466847897
step: 260, loss: 0.005317769479006529
step: 270, loss: 0.0048530991189181805
step: 280, loss: 0.007204296067357063
step: 290, loss: 0.001268828404136002
step: 300, loss: 0.0005113172228448093
step: 310, loss: 0.0035014781169593334
step: 320, loss: 0.030270343646407127
step: 330, loss: 0.03656366094946861
step: 340, loss: 0.0007633990026079118
step: 350, loss: 0.0286440160125494
step: 360, loss: 0.0008053990313783288
step: 370, loss: 0.010982231236994267
step: 380, loss: 0.006619411054998636
step: 390, loss: 0.0020539502147585154

```

```

step: 400, loss: 0.006364364642649889
step: 410, loss: 0.004089065361768007
step: 420, loss: 0.02141963317990303
step: 430, loss: 0.0001789464004104957
step: 440, loss: 0.03491724655032158
step: 450, loss: 0.0009617909672670066
step: 460, loss: 0.0020314622670412064
step: 470, loss: 0.016041390597820282
step: 480, loss: 0.019016744568943977
step: 490, loss: 0.000276533595751971
step: 500, loss: 0.0028495099395513535
step: 510, loss: 0.0013454967411234975
step: 520, loss: 0.0012483577011153102
step: 530, loss: 0.002080460311844945
/usr/local/lib/python3.7/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=0.9075785582255084, f1=0.902867715078631,
best_f1=0.9138712601994561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004993367474526167
step: 10, loss: 0.012438981793820858
step: 20, loss: 0.012166996486485004
step: 30, loss: 0.013015282340347767
step: 40, loss: 0.0006208601989783347
step: 50, loss: 0.0009116960573010147
step: 60, loss: 0.00395544245839119
step: 70, loss: 0.0007837900775484741
step: 80, loss: 0.0003991495177615434
step: 90, loss: 0.0005720345070585608
step: 100, loss: 0.0003942592884413898
step: 110, loss: 0.0004353425174485892
step: 120, loss: 0.000182490810402669
step: 130, loss: 0.0002508870675228536
step: 140, loss: 0.0006379150436259806
step: 150, loss: 0.00011436678323661909
step: 160, loss: 0.00028184076654724777
step: 170, loss: 0.019440406933426857
step: 180, loss: 0.00019935620366595685
step: 190, loss: 0.0001618833193788305
step: 200, loss: 0.00015293936303351074
step: 210, loss: 0.00022681835980620235
step: 220, loss: 0.00016782659804448485
step: 230, loss: 0.000959097349550575
step: 240, loss: 0.007877101190388203
step: 250, loss: 0.008455253206193447

```

```

step: 260, loss: 0.00022517687466461211
step: 270, loss: 0.00010998345533153042
step: 280, loss: 0.0031563411466777325
step: 290, loss: 0.025959717109799385
step: 300, loss: 0.0007034888840280473
step: 310, loss: 0.0010774858528748155
step: 320, loss: 0.007198908366262913
step: 330, loss: 0.0009424573509022593
step: 340, loss: 0.0011522972490638494
step: 350, loss: 0.0012546968646347523
step: 360, loss: 0.0018926870543509722
step: 370, loss: 0.0009441970032639802
step: 380, loss: 0.002083495492115617
step: 390, loss: 0.0005402888054959476
step: 400, loss: 0.0002880140964407474
step: 410, loss: 0.014215008355677128
step: 420, loss: 0.002140172990038991
step: 430, loss: 0.00022536581673193723
step: 440, loss: 0.0021949338261038065
step: 450, loss: 0.0024758547078818083
step: 460, loss: 0.0018303532851859927
step: 470, loss: 0.06243418902158737
step: 480, loss: 0.0008610254735685885
step: 490, loss: 0.00023321215121541172
step: 500, loss: 0.0001171850526588969
step: 510, loss: 0.05118891969323158
step: 520, loss: 0.002615272533148527
step: 530, loss: 0.0003199049679096788
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9184909175593852, f1=0.9126213592233011,
best_f1=0.9126213592233011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028087340760976076
step: 10, loss: 0.0014335380401462317
step: 20, loss: 0.002077485201880336
step: 30, loss: 0.007886754348874092
step: 40, loss: 0.0004225435550324619
step: 50, loss: 0.0003063002077396959
step: 60, loss: 0.02236887440085411
step: 70, loss: 0.0002114053932018578

```

```
step: 80, loss: 0.0007210209150798619
step: 90, loss: 0.0012442150618880987
step: 100, loss: 0.0015036421827971935
step: 110, loss: 0.0027806947473436594
step: 120, loss: 0.002385294996201992
step: 130, loss: 0.00081073452020064
step: 140, loss: 6.299384403973818e-05
step: 150, loss: 0.0008483737474307418
step: 160, loss: 0.00011558826372493058
step: 170, loss: 0.004836961627006531
step: 180, loss: 0.0009973758133128285
step: 190, loss: 0.0011855245102196932
step: 200, loss: 0.00049973075510934
step: 210, loss: 0.0002809331053867936
step: 220, loss: 0.00039265997475013137
step: 230, loss: 0.00048687076196074486
step: 240, loss: 0.00020133111684117466
step: 250, loss: 0.0014047448057681322
step: 260, loss: 0.026615696027874947
step: 270, loss: 0.00028929501422680914
step: 280, loss: 0.00023637004778720438
step: 290, loss: 0.0006946963258087635
step: 300, loss: 0.00010661785199772567
step: 310, loss: 0.060053501278162
step: 320, loss: 9.789039904717356e-05
step: 330, loss: 0.0001390163815813139
step: 340, loss: 0.012817390263080597
step: 350, loss: 0.00035005484824068844
step: 360, loss: 0.00014182312588673085
step: 370, loss: 0.00013119718641974032
step: 380, loss: 0.10049550980329514
step: 390, loss: 8.616691775387153e-05
step: 400, loss: 0.033847782760858536
step: 410, loss: 0.00012767604494001716
step: 420, loss: 0.000139591284096241
step: 430, loss: 0.09744956344366074
step: 440, loss: 0.05317221209406853
step: 450, loss: 0.0005832815077155828
step: 460, loss: 0.0013222111156210303
step: 470, loss: 0.1214989498257637
step: 480, loss: 0.0038075819611549377
step: 490, loss: 0.00017979023687075824
step: 500, loss: 0.0021193900611251593
step: 510, loss: 0.0007320625009015203
step: 520, loss: 0.013352741487324238
step: 530, loss: 0.0006356380763463676
/usr/local/lib/python3.7/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=0.914844113541182, f1=0.916937354988399,
best_f1=0.9126213592233011
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019776911358349025
step: 10, loss: 0.07494539767503738
step: 20, loss: 0.0002205859200330451
step: 30, loss: 0.0030672254506498575
step: 40, loss: 0.0012866408796980977
step: 50, loss: 0.0006593117141164839
step: 60, loss: 0.01055878959596157
step: 70, loss: 0.025592762976884842
step: 80, loss: 0.000739753304515034
step: 90, loss: 0.001096022198908031
step: 100, loss: 0.026827406138181686
step: 110, loss: 0.0015934883849695325
step: 120, loss: 0.0001503134408267215
step: 130, loss: 0.0005198205471970141
step: 140, loss: 0.05037578567862511
step: 150, loss: 0.0006426603649742901
step: 160, loss: 0.00010962517990265042
step: 170, loss: 0.0018093526596203446
step: 180, loss: 0.00013978334027342498
step: 190, loss: 0.0018237752374261618
step: 200, loss: 0.003434631507843733
step: 210, loss: 5.624128243653104e-05
step: 220, loss: 0.0294246356934309
step: 230, loss: 9.287210559705272e-05
step: 240, loss: 0.00014315707085188478
step: 250, loss: 0.00034486170625314116
step: 260, loss: 0.00046427708002738655
step: 270, loss: 0.028948364779353142
step: 280, loss: 7.156370702432469e-05
step: 290, loss: 6.394201045623049e-05
step: 300, loss: 0.013069548644125462
step: 310, loss: 0.006432565860450268
step: 320, loss: 0.00016120057262014598
step: 330, loss: 0.00013263107393868268
step: 340, loss: 0.00031878502340987325
step: 350, loss: 0.000267862225882709
step: 360, loss: 0.00357546447776258
step: 370, loss: 0.00021488085621967912
step: 380, loss: 5.186221460462548e-05
step: 390, loss: 0.0007458429899998009
step: 400, loss: 0.005462768953293562
step: 410, loss: 0.00026925475685857236

```

```

step: 420, loss: 0.010504160076379776
step: 430, loss: 5.7733621360966936e-05
step: 440, loss: 0.004939081612974405
step: 450, loss: 0.00010738630953710526
step: 460, loss: 6.921065505594015e-05
step: 470, loss: 0.00010819335147971287
step: 480, loss: 0.00018037004338111728
step: 490, loss: 8.843989780871198e-05
step: 500, loss: 0.0002357645716983825
step: 510, loss: 6.390500493580475e-05
step: 520, loss: 6.992024282226339e-05
step: 530, loss: 0.005013743415474892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9235048678720444, f1=0.917397323488694,
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007166742580011487
step: 10, loss: 0.00011562916188267991
step: 20, loss: 0.00013844112982042134
step: 30, loss: 4.4335374695947394e-05
step: 40, loss: 0.001752569922246039
step: 50, loss: 5.193702600081451e-05
step: 60, loss: 6.362043495755643e-05
step: 70, loss: 7.807343354215845e-05
step: 80, loss: 0.0002753855078481138
step: 90, loss: 4.651887866202742e-05
step: 100, loss: 0.0001984818809432909
step: 110, loss: 0.00010472015856066719
step: 120, loss: 0.00111136375926435
step: 130, loss: 0.02391224168241024
step: 140, loss: 7.029308471828699e-05
step: 150, loss: 0.00029926904244348407
step: 160, loss: 3.2498410291736946e-05
step: 170, loss: 0.0062964013777673244
step: 180, loss: 5.2253850299166515e-05
step: 190, loss: 0.00029250700026750565
step: 200, loss: 0.00803154893219471
step: 210, loss: 4.330827869125642e-05
step: 220, loss: 0.00011351007560733706
step: 230, loss: 9.680774383014068e-05

```

```

step: 240, loss: 0.00018568853556644171
step: 250, loss: 0.00022463970526587218
step: 260, loss: 0.0010631612967699766
step: 270, loss: 5.2001232688780874e-05
step: 280, loss: 0.0007491987198591232
step: 290, loss: 8.45196918817237e-05
step: 300, loss: 0.0004931987496092916
step: 310, loss: 6.309252057690173e-05
step: 320, loss: 3.9195332647068426e-05
step: 330, loss: 0.0005093667423352599
step: 340, loss: 0.000464678363641724
step: 350, loss: 0.000384705577744171
step: 360, loss: 0.0001427783427061513
step: 370, loss: 0.00032684230245649815
step: 380, loss: 0.0002771181461866945
step: 390, loss: 2.9965418434585445e-05
step: 400, loss: 8.379250357393175e-05
step: 410, loss: 4.8405592679046094e-05
step: 420, loss: 0.0025127020198851824
step: 430, loss: 9.497132850810885e-05
step: 440, loss: 4.4797245209338143e-05
step: 450, loss: 0.0016258410178124905
step: 460, loss: 0.00014034607738722116
step: 470, loss: 3.433434903854504e-05
step: 480, loss: 0.0004969018045812845
step: 490, loss: 0.0698518455028534
step: 500, loss: 0.006066733971238136
step: 510, loss: 0.000206983313546516
step: 520, loss: 6.666366971330717e-05
step: 530, loss: 0.005004099104553461
/usr/local/lib/python3.7/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=0.912067352666043, f1=0.912901723334886,
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005894128116779029
step: 10, loss: 0.004892267752438784
step: 20, loss: 0.0010223384015262127
step: 30, loss: 0.0008177530253306031
step: 40, loss: 0.01335901115089655
step: 50, loss: 0.0003142307105008513
step: 60, loss: 0.00020017231872770935
step: 70, loss: 7.686368189752102e-05
step: 80, loss: 0.0005836659693159163
step: 90, loss: 0.0003195804893039167

```

```
step: 100, loss: 3.225575346732512e-05
step: 110, loss: 8.103041182039306e-05
step: 120, loss: 0.0001412459387211129
step: 130, loss: 4.068817725055851e-05
step: 140, loss: 0.1553427278995514
step: 150, loss: 0.0021742971148341894
step: 160, loss: 0.0001174598146462813
step: 170, loss: 0.0011142587754875422
step: 180, loss: 6.070323797757737e-05
step: 190, loss: 0.0007675504311919212
step: 200, loss: 0.0009839143604040146
step: 210, loss: 6.475616828538477e-05
step: 220, loss: 0.0003225442487746477
step: 230, loss: 8.976247045211494e-05
step: 240, loss: 5.574548049480654e-05
step: 250, loss: 0.0001697016996331513
step: 260, loss: 7.952483429107815e-05
step: 270, loss: 0.000633072224445641
step: 280, loss: 0.0014031725004315376
step: 290, loss: 0.017633141949772835
step: 300, loss: 0.012535862624645233
step: 310, loss: 0.006097863428294659
step: 320, loss: 0.0028160754591226578
step: 330, loss: 0.007110972888767719
step: 340, loss: 4.226093733450398e-05
step: 350, loss: 0.23077461123466492
step: 360, loss: 0.00018065079348161817
step: 370, loss: 4.981319943908602e-05
step: 380, loss: 0.00013635742652695626
step: 390, loss: 0.0007027317769825459
step: 400, loss: 5.96297177253291e-05
step: 410, loss: 0.0016375657869502902
step: 420, loss: 0.0026279822923243046
step: 430, loss: 0.02492869272828102
step: 440, loss: 9.693064930615947e-05
step: 450, loss: 0.00020008947467431426
step: 460, loss: 0.0018302815733477473
step: 470, loss: 0.010348507203161716
step: 480, loss: 0.0003833655791822821
step: 490, loss: 8.905029972083867e-05
step: 500, loss: 9.625960956327617e-05
step: 510, loss: 4.359884042060003e-05
step: 520, loss: 6.888576172059402e-05
step: 530, loss: 0.0001665228046476841
/usr/local/lib/python3.7/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=0.9177449168207023, f1=0.9152073732718895,
```

```
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013594968186225742
step: 10, loss: 0.00013128969294484705
step: 20, loss: 8.338898624060676e-05
step: 30, loss: 6.353231583489105e-05
step: 40, loss: 7.979846122907475e-05
step: 50, loss: 0.015825672075152397
step: 60, loss: 3.786555680562742e-05
step: 70, loss: 0.000102759622677695
step: 80, loss: 0.0038636885583400726
step: 90, loss: 4.983710095984861e-05
step: 100, loss: 0.0002841036766767502
step: 110, loss: 0.0004909816780127585
step: 120, loss: 5.246502041700296e-05
step: 130, loss: 8.200219599530101e-05
step: 140, loss: 3.609309715102427e-05
step: 150, loss: 0.001245745806954801
step: 160, loss: 0.00021262431982904673
step: 170, loss: 0.00019907348905690014
step: 180, loss: 5.3684307204093784e-05
step: 190, loss: 4.0878581785364076e-05
step: 200, loss: 7.844207721063867e-05
step: 210, loss: 8.410769078182057e-05
step: 220, loss: 3.376810491317883e-05
step: 230, loss: 0.00015190953854471445
step: 240, loss: 0.00011266415094723925
step: 250, loss: 0.0008920427062548697
step: 260, loss: 0.00011351393914083019
step: 270, loss: 5.325241363607347e-05
step: 280, loss: 6.680205115117133e-05
step: 290, loss: 0.003299225587397814
step: 300, loss: 0.018790986388921738
step: 310, loss: 0.00013204575225245208
step: 320, loss: 5.5457421694882214e-05
step: 330, loss: 2.986471554322634e-05
step: 340, loss: 2.716787093959283e-05
step: 350, loss: 0.0013866339577361941
step: 360, loss: 5.4047217417974025e-05
step: 370, loss: 2.8496851882664487e-05
step: 380, loss: 0.0001146171271102503
step: 390, loss: 0.00037588385748676956
step: 400, loss: 4.5627053623320535e-05
step: 410, loss: 0.0001289415522478521
step: 420, loss: 7.80898699304089e-05
step: 430, loss: 5.949096521362662e-05
```

```

step: 440, loss: 0.0014195096446201205
step: 450, loss: 0.0034969495609402657
step: 460, loss: 9.842324652709067e-05
step: 470, loss: 4.462936340132728e-05
step: 480, loss: 0.000522034359164536
step: 490, loss: 0.00023318734019994736
step: 500, loss: 8.48364143166691e-05
step: 510, loss: 3.9855160139268264e-05
step: 520, loss: 0.001190945622511208
step: 530, loss: 3.895717964041978e-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=0.9154734411085451, f1=0.9143914854234151,
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017461585230194032
step: 10, loss: 1.842132769525051e-05
step: 20, loss: 0.0006300446693785489
step: 30, loss: 0.0002450913016218692
step: 40, loss: 3.1190782465273514e-05
step: 50, loss: 9.355024667456746e-05
step: 60, loss: 3.1715884688310325e-05
step: 70, loss: 8.437321957899258e-05
step: 80, loss: 5.012533802073449e-05
step: 90, loss: 0.0001820049510570243
step: 100, loss: 0.00662140641361475
step: 110, loss: 4.9692542233970016e-05
step: 120, loss: 0.0008086543530225754
step: 130, loss: 0.003912786487489939
step: 140, loss: 6.880805449327454e-05
step: 150, loss: 4.364882624940947e-05
step: 160, loss: 2.9246371923363768e-05
step: 170, loss: 3.840236604446545e-05
step: 180, loss: 5.167997733224183e-05
step: 190, loss: 0.00011409545550122857
step: 200, loss: 0.00017306621884927154
step: 210, loss: 7.503638335037977e-05
step: 220, loss: 1.981824971153401e-05
step: 230, loss: 4.227282988722436e-05
step: 240, loss: 4.290550350560807e-05
step: 250, loss: 0.004324465524405241
step: 260, loss: 0.00023499989765696228
step: 270, loss: 0.00019294509547762573
step: 280, loss: 0.00016201821563299745
step: 290, loss: 8.914920181268826e-05

```

```

step: 300, loss: 4.3220781662967056e-05
step: 310, loss: 2.1975041818222962e-05
step: 320, loss: 2.4273531380458735e-05
step: 330, loss: 0.00016576770576648414
step: 340, loss: 2.542083711887244e-05
step: 350, loss: 5.925508594373241e-05
step: 360, loss: 4.5858643716201186e-05
step: 370, loss: 0.0006712925969623029
step: 380, loss: 4.4763542973669246e-05
step: 390, loss: 3.189119161106646e-05
step: 400, loss: 2.798652531055268e-05
step: 410, loss: 0.0002357369230594486
step: 420, loss: 2.475778092048131e-05
step: 430, loss: 5.2300747483968735e-05
step: 440, loss: 0.0004481514624785632
step: 450, loss: 3.99778873543255e-05
step: 460, loss: 7.638301030965522e-05
step: 470, loss: 0.0007547312416136265
step: 480, loss: 3.884631587425247e-05
step: 490, loss: 4.128110231249593e-05
step: 500, loss: 0.010274260304868221
step: 510, loss: 0.0001397905871272087
step: 520, loss: 3.6693920264951885e-05
step: 530, loss: 8.729177352506667e-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=0.9163568773234201, f1=0.9183202584217812,
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011037926014978439
step: 10, loss: 0.00025776113034226
step: 20, loss: 8.165874896803871e-05
step: 30, loss: 6.822594878030941e-05
step: 40, loss: 2.983392187161371e-05
step: 50, loss: 3.948575977119617e-05
step: 60, loss: 2.9458784410962835e-05
step: 70, loss: 2.3919428713270463e-05
step: 80, loss: 5.612920722342096e-05
step: 90, loss: 2.3107400920707732e-05
step: 100, loss: 3.0624738428741693e-05
step: 110, loss: 2.2857888325233944e-05
step: 120, loss: 2.2939762857276946e-05
step: 130, loss: 3.4615623007994145e-05
step: 140, loss: 0.0022890837863087654
step: 150, loss: 2.6288869776180945e-05

```

```

step: 160, loss: 8.262924529844895e-05
step: 170, loss: 4.889623596682213e-05
step: 180, loss: 0.030188705772161484
step: 190, loss: 0.00017022871179506183
step: 200, loss: 5.1566381443990394e-05
step: 210, loss: 0.00019667288870550692
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.0003144282381981611
step: 230, loss: 4.36863956565503e-05
step: 240, loss: 2.353584932279773e-05
step: 250, loss: 6.477045826613903e-05
step: 260, loss: 1.8551612811279483e-05
step: 270, loss: 0.002938241930678487
step: 280, loss: 2.1017776816734113e-05
step: 290, loss: 5.2884093747707084e-05
step: 300, loss: 0.00013533799210563302
step: 310, loss: 2.9596476451843046e-05
step: 320, loss: 0.002613854594528675
step: 330, loss: 7.526385161327198e-05
step: 340, loss: 0.0004072637821082026
step: 350, loss: 0.0008343373774550855
step: 360, loss: 0.001398711814545095
step: 370, loss: 9.5887909992598e-05
step: 380, loss: 4.360879393061623e-05
step: 390, loss: 4.526770135271363e-05
step: 400, loss: 2.57785231951857e-05
step: 410, loss: 2.511535058147274e-05
step: 420, loss: 2.5145049221464433e-05
step: 430, loss: 3.0352613975992426e-05
step: 440, loss: 2.5803268727031536e-05
step: 450, loss: 0.0001375876454403624
step: 460, loss: 0.00011580788850551471
step: 470, loss: 4.625607834896073e-05
step: 480, loss: 0.00012886623153463006
step: 490, loss: 2.678355303942226e-05
step: 500, loss: 8.952784992288798e-05
step: 510, loss: 6.151564593892545e-05
step: 520, loss: 3.294476482551545e-05
step: 530, loss: 2.3662571038585156e-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=0.916628281897743, f1=0.9192660550458716,
best_f1=0.917397323488694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022717851970810443

```


step: 10, loss: 5.01426839036867e-05
step: 20, loss: 6.769929314032197e-05
step: 30, loss: 4.660032573156059e-05
step: 40, loss: 2.6030573280877434e-05
step: 50, loss: 7.268938497873023e-05
step: 60, loss: 1.8220143829239532e-05
step: 70, loss: 2.4772600227151997e-05
step: 80, loss: 2.776745350274723e-05
step: 90, loss: 2.3929924282128923e-05
step: 100, loss: 2.55323393503204e-05
step: 110, loss: 0.0005042359116487205
step: 120, loss: 0.020372137427330017
step: 130, loss: 2.290246447955724e-05
step: 140, loss: 1.904341479530558e-05
step: 150, loss: 0.000773193605709821
step: 160, loss: 0.0001510578440502286
step: 170, loss: 2.2537506083608605e-05
step: 180, loss: 5.529026020667516e-05
step: 190, loss: 0.0002280316111864522
step: 200, loss: 2.2250702386372723e-05
step: 210, loss: 3.123532951576635e-05
step: 220, loss: 2.2172114768181928e-05
step: 230, loss: 5.108220284455456e-05
step: 240, loss: 8.03618022473529e-05
step: 250, loss: 7.732411177130416e-05
step: 260, loss: 7.45600918889977e-05
step: 270, loss: 0.0012848295737057924
step: 280, loss: 0.00011756443564081565
step: 290, loss: 0.00028924059006385505
step: 300, loss: 4.062479638378136e-05
step: 310, loss: 0.00013013285933993757
step: 320, loss: 7.407831435557455e-05
step: 330, loss: 3.880770964315161e-05
step: 340, loss: 8.268133387900889e-05
step: 350, loss: 3.0973886168794706e-05
step: 360, loss: 3.197591286152601e-05
step: 370, loss: 2.0362092982395552e-05
step: 380, loss: 3.487250069156289e-05
step: 390, loss: 0.003702036803588271
step: 400, loss: 0.0005341147189028561
step: 410, loss: 7.433399878209457e-05
step: 420, loss: 6.032617238815874e-05
step: 430, loss: 2.3066455469233915e-05
step: 440, loss: 8.112466457532719e-05
step: 450, loss: 5.4944390285527334e-05
step: 460, loss: 1.8462184016243555e-05
step: 470, loss: 3.0231070923036896e-05
step: 480, loss: 4.244866067892872e-05

```

step: 490, loss: 3.331723564770073e-05
step: 500, loss: 3.9354112232103944e-05
step: 510, loss: 0.00022715561499353498
step: 520, loss: 2.815435618686024e-05
step: 530, loss: 4.3522424675757065e-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=0.9172445677300047, f1=0.9194661757938334,
best_f1=0.917397323488694
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 [00:16, 341.61it/s]
load_f1 = 0.9228624535315985
real_f1 = 0.9229340761374188
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 [00:12, 363.31it/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 1.2 MB/s

Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)

Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)

Requirement already satisfied: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)

Requirement already satisfied: 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: 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: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)

Requirement already satisfied: plac<1.2.0,>=0.9.6 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.1.3)

Requirement already satisfied: 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: 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: 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: 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: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)

Requirement already satisfied: wasabi<1.1.0,>=0.4.0 in

```

/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: 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: 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: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: 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=8de6900bbb8b950816af3560bc7b57b8a84b6a55e80fc5c799f3bd990de4a7f0
  Stored in directory: /tmp/pip-ephem-wheel-cache-018eyipo/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 distilbert \
  --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 distilbert \
  --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: 0.8715479969978333
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=0.3333333333333333, f1=0.15789473684210525,
best_f1=0.15789473684210525
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3770557940006256
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37209302325581395, f1=0.1935483870967742,
best_f1=0.1935483870967742

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33033618330955505
/usr/local/lib/python3.7/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=0.35135135135135137, f1=0.3880597014925373,
best_f1=0.1935483870967742
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3558439314365387
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.43478260869565216, f1=0.36363636363636365,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2722797095775604
/usr/local/lib/python3.7/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=0.42622950819672134, f1=0.44444444444444444,
best_f1=0.36363636363636365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3108725845813751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5217391304347825, f1=0.25, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21567930281162262
/usr/local/lib/python3.7/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=0.423076923076923, f1=0.3829787234042553, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3573867678642273
/usr/local/lib/python3.7/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=0.40816326530612246, f1=0.4782608695652174, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17977721989154816
/usr/local/lib/python3.7/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=0.44897959183673464, f1=0.43478260869565216, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2498604655265808
/usr/local/lib/python3.7/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=0.45, f1=0.3529411764705882, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25528794527053833
/usr/local/lib/python3.7/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=0.4782608695652174, f1=0.5365853658536585, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29203352332115173
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5238095238095237, f1=0.4117647058823529,
best_f1=0.4117647058823529

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22232933342456818
/usr/local/lib/python3.7/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=0.5, f1=0.23999999999999996, best_f1=0.4117647058823529
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24082306027412415
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5384615384615384, f1=0.23076923076923075,
best_f1=0.23076923076923075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25761696696281433
/usr/local/lib/python3.7/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=0.5384615384615384, f1=0.23076923076923075,
best_f1=0.23076923076923075
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, 141573.32it/s]
/usr/local/lib/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 = 0.6153846153846153
real_f1 = 0.5714285714285714
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 [00:10, 422.65it/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 distilbert \
    --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 distilbert \
--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: 0.7898659110069275

step: 10, loss: 0.4707822799682617

step: 20, loss: 0.5995522141456604

```

step: 30, loss: 0.46988558769226074
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.391080379486084
step: 50, loss: 0.19137118756771088
step: 60, loss: 0.20342154800891876
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 70, loss: 0.3443858325481415
step: 80, loss: 0.2087552696466446
step: 90, loss: 0.058589253574609756
step: 100, loss: 0.08101368695497513
step: 110, loss: 0.040984489023685455
step: 120, loss: 0.030312327668070793
step: 130, loss: 0.21951431035995483
step: 140, loss: 0.1298820525407791
step: 150, loss: 0.0501391626894474
step: 160, loss: 0.1516437530517578
step: 170, loss: 0.006819815840572119
step: 180, loss: 0.027299147099256516
step: 190, loss: 0.055860504508018494
step: 200, loss: 0.04638348147273064
step: 210, loss: 0.057972900569438934
step: 220, loss: 0.0064042857848107815
step: 230, loss: 0.023121792823076248
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9796839729119639, f1=0.9695603156708005,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022875722497701645
step: 10, loss: 0.051897380501031876
step: 20, loss: 0.009364455007016659
step: 30, loss: 0.00261243456043303
step: 40, loss: 0.04364289715886116
step: 50, loss: 0.0055551384575665
step: 60, loss: 0.003032352775335312
step: 70, loss: 0.27267318964004517
step: 80, loss: 0.0034615565091371536
step: 90, loss: 0.03622037172317505
step: 100, loss: 0.0661105290055275
step: 110, loss: 0.008368736132979393
step: 120, loss: 0.0291207917034626

```

```

step: 130, loss: 0.002833673032000661
step: 140, loss: 0.11288147419691086
step: 150, loss: 0.007724879775196314
step: 160, loss: 0.047833144664764404
step: 170, loss: 0.20594103634357452
step: 180, loss: 0.003744882997125387
step: 190, loss: 0.23622044920921326
step: 200, loss: 0.005644222721457481
step: 210, loss: 0.03107515349984169
step: 220, loss: 0.0016984494868665934
step: 230, loss: 0.04776301980018616
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.967305524239008,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1283624917268753
step: 10, loss: 0.018331162631511688
step: 20, loss: 0.00731821171939373
step: 30, loss: 0.014507138170301914
step: 40, loss: 0.011607854627072811
step: 50, loss: 0.03160151094198227
step: 60, loss: 0.004032148513942957
step: 70, loss: 0.04608684778213501
step: 80, loss: 0.017899751663208008
step: 90, loss: 0.0036496049724519253
step: 100, loss: 0.003839637851342559
step: 110, loss: 0.03936881944537163
step: 120, loss: 0.0018849298357963562
step: 130, loss: 0.0003685243136715144
step: 140, loss: 0.02241024561226368
step: 150, loss: 0.002756670117378235
step: 160, loss: 0.000987983657978475
step: 170, loss: 0.0012426543980836868
step: 180, loss: 0.0018493399256840348
step: 190, loss: 0.004580571781843901
step: 200, loss: 0.00033126637572422624
step: 210, loss: 0.16342663764953613
step: 220, loss: 0.009181229397654533
step: 230, loss: 0.008908547461032867
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9720670391061451,
best_f1=0.9695603156708005

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009399693459272385
step: 10, loss: 0.013080613687634468
step: 20, loss: 0.0009956156136468053
step: 30, loss: 0.00035901236697100103
step: 40, loss: 0.010471970774233341
step: 50, loss: 0.0012133745476603508
step: 60, loss: 0.0006155883893370628
step: 70, loss: 0.0924648568034172
step: 80, loss: 0.10709139704704285
step: 90, loss: 0.02060881070792675
step: 100, loss: 0.009164799936115742
step: 110, loss: 0.0037533061113208532
step: 120, loss: 0.006508713588118553
step: 130, loss: 0.00173191970679909
step: 140, loss: 0.0031269409228116274
step: 150, loss: 0.0011825429974123836
step: 160, loss: 0.0010437627788633108
step: 170, loss: 0.002836005762219429
step: 180, loss: 0.034132689237594604
step: 190, loss: 0.006331481039524078
step: 200, loss: 0.002617072081193328
step: 210, loss: 0.001426179544068873
step: 220, loss: 0.0006984042120166123
step: 230, loss: 0.0007858058670535684
/usr/local/lib/python3.7/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=0.977728285077951, f1=0.9775784753363228,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039641448529437184
step: 10, loss: 0.001121815643273294
step: 20, loss: 0.00036149602965451777
step: 30, loss: 0.013608451001346111
step: 40, loss: 0.00019276402599643916
step: 50, loss: 0.00012706994311884046
step: 60, loss: 0.023787422105669975
step: 70, loss: 0.004335578065365553
step: 80, loss: 0.009059052914381027
step: 90, loss: 0.03869664669036865
step: 100, loss: 0.003592191729694605
step: 110, loss: 0.0012574357679113746
step: 120, loss: 0.08291704952716827

```

```

step: 130, loss: 0.006268157623708248
step: 140, loss: 0.0025743735022842884
step: 150, loss: 0.0003606012905947864
step: 160, loss: 0.0047266511246562
step: 170, loss: 0.0007855738513171673
step: 180, loss: 0.0012929975055158138
step: 190, loss: 0.0001278318522963673
step: 200, loss: 0.0035361158661544323
step: 210, loss: 7.584327977383509e-05
step: 220, loss: 0.0002364717365708202
step: 230, loss: 0.003045669523999095
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842342342342343, f1=0.9786276715410572,
best_f1=0.9786276715410572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002946085878647864
step: 10, loss: 0.0004003283102065325
step: 20, loss: 0.00012084538320777938
step: 30, loss: 0.00011186279880348593
step: 40, loss: 0.00034127419348806143
step: 50, loss: 0.0047665247693657875
step: 60, loss: 0.011259709484875202
step: 70, loss: 0.00017102509445976466
step: 80, loss: 0.006512379739433527
step: 90, loss: 9.848461195360869e-05
step: 100, loss: 0.0003140653425361961
step: 110, loss: 0.026132192462682724
step: 120, loss: 0.0005691344849765301
step: 130, loss: 0.005842362996190786
step: 140, loss: 0.0008781535434536636
step: 150, loss: 0.0002812245220411569
step: 160, loss: 0.017251739278435707
step: 170, loss: 0.00046440717414952815
step: 180, loss: 0.000664838938973844
step: 190, loss: 0.005108379293233156
step: 200, loss: 0.0011470766039565206
step: 210, loss: 0.0021356921643018723
step: 220, loss: 0.0007870005210861564
step: 230, loss: 0.00018022386939264834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 6: dev_f1=0.9853768278965129, f1=0.9775280898876404,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008484047139063478
step: 10, loss: 0.00038207985926419497
step: 20, loss: 0.0003851840447168797
step: 30, loss: 0.00015959526353981346
step: 40, loss: 0.00039803236722946167
step: 50, loss: 0.00016047283133957535
step: 60, loss: 0.00010947990813292563
step: 70, loss: 0.00011840023216791451
step: 80, loss: 6.509669765364379e-05
step: 90, loss: 0.0007725388277322054
step: 100, loss: 0.00032628607004880905
step: 110, loss: 0.0055701471865177155
step: 120, loss: 0.0007461664499714971
step: 130, loss: 0.0005188641953282058
step: 140, loss: 0.00022687834280077368
step: 150, loss: 0.00020229924120940268
step: 160, loss: 0.00017776071035768837
step: 170, loss: 0.19108368456363678
step: 180, loss: 0.0018228425178676844
step: 190, loss: 0.0005254152929410338
step: 200, loss: 0.004358852282166481
step: 210, loss: 7.435290171997622e-05
step: 220, loss: 0.000313945027301088
step: 230, loss: 0.07524411380290985
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9843400447427293, f1=0.9743016759776536,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07516900449991226
step: 10, loss: 0.00041390021215192974
step: 20, loss: 0.00039247312815859914
step: 30, loss: 0.0003405030001886189
step: 40, loss: 0.00036672805435955524

```



```

step: 50, loss: 0.0027252272702753544
step: 60, loss: 0.04045693576335907
step: 70, loss: 0.0021791374310851097
step: 80, loss: 0.0015927812783047557
step: 90, loss: 0.0019158001523464918
step: 100, loss: 0.0009225911926478148
step: 110, loss: 0.0009744000853970647
step: 120, loss: 0.0003001336590386927
step: 130, loss: 0.019071871414780617
step: 140, loss: 0.0003478246508166194
step: 150, loss: 0.0017282450571656227
step: 160, loss: 0.00030853322823531926
step: 170, loss: 0.12529931962490082
step: 180, loss: 0.001634174957871437
step: 190, loss: 0.0003520379541441798
step: 200, loss: 0.004633962642401457
step: 210, loss: 0.005205857567489147
step: 220, loss: 0.00030154158594086766
step: 230, loss: 0.00090078916400671
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9865168539325843, 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: 0.0006144627113826573
step: 10, loss: 0.0002448717423249036
step: 20, loss: 0.0007985320407897234
step: 30, loss: 0.0003438717103563249
step: 40, loss: 0.00040690210880711675
step: 50, loss: 0.000563062378205359
step: 60, loss: 0.0019076191820204258
step: 70, loss: 0.003116637235507369
step: 80, loss: 0.0006671531009487808
step: 90, loss: 0.0008620696607977152
step: 100, loss: 0.000365958345355466
step: 110, loss: 0.00027666633832268417
step: 120, loss: 0.0015258094063028693
step: 130, loss: 0.00017168832710012794
step: 140, loss: 0.00019074765441473573
step: 150, loss: 0.00045615751878358424
step: 160, loss: 0.00042534267413429916

```

```

step: 170, loss: 9.017373668029904e-05
step: 180, loss: 0.00023785338271409273
step: 190, loss: 0.0003268971049692482
step: 200, loss: 0.0001379455061396584
step: 210, loss: 0.0006538642919622362
step: 220, loss: 0.09973789751529694
step: 230, loss: 0.0009648095001466572
/usr/local/lib/python3.7/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=0.9774774774774775,
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: 0.0004394257557578385
step: 10, loss: 8.52529046824202e-05
step: 20, loss: 0.0001696896506473422
step: 30, loss: 0.000604424043558538
step: 40, loss: 0.00015915892436169088
step: 50, loss: 0.013095003552734852
step: 60, loss: 0.0005331370630301535
step: 70, loss: 0.0006106992950662971
step: 80, loss: 0.0004306703922338784
step: 90, loss: 0.00010689473856473342
step: 100, loss: 0.0001250601198989898
step: 110, loss: 0.00016008013335522264
step: 120, loss: 0.0034529350232332945
step: 130, loss: 0.00019898783648386598
step: 140, loss: 0.0011965656885877252
step: 150, loss: 6.926345668034628e-05
step: 160, loss: 0.0003262285317759961
step: 170, loss: 0.00027136417338624597
step: 180, loss: 0.00015097358846105635
step: 190, loss: 0.00010119377839146182
step: 200, loss: 7.738585554761812e-05
step: 210, loss: 9.313285409007221e-05
step: 220, loss: 0.0019098774064332247
step: 230, loss: 8.31045545055531e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9832402234636871, f1=0.9798657718120806,
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: 0.0010284233139827847

```

```

step: 10, loss: 0.0041184984147548676
step: 20, loss: 4.8079506086651236e-05
step: 30, loss: 0.0001572353212395683
step: 40, loss: 0.0005832146271131933
step: 50, loss: 0.005703567061573267
step: 60, loss: 0.0001411245611961931
step: 70, loss: 0.00019428555970080197
step: 80, loss: 0.00017182949522975832
step: 90, loss: 0.0004281600995454937
step: 100, loss: 0.0001561171084176749
step: 110, loss: 0.0001316428097197786
step: 120, loss: 0.008584690280258656
step: 130, loss: 6.295941420830786e-05
step: 140, loss: 0.0015911714872345328
step: 150, loss: 8.138232806231827e-05
step: 160, loss: 0.0001504750398453325
step: 170, loss: 0.0005658887675963342
step: 180, loss: 0.008092333562672138
step: 190, loss: 0.0001663079601712525
step: 200, loss: 6.562835187651217e-05
step: 210, loss: 5.4260046454146504e-05
step: 220, loss: 9.927947394317016e-05
step: 230, loss: 0.00045715196756646037
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9787234042553192,
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: 0.00010404564818600193
step: 10, loss: 0.0026970882900059223
step: 20, loss: 4.601279579219408e-05
step: 30, loss: 0.061270229518413544
step: 40, loss: 4.9106907681562006e-05
step: 50, loss: 4.437704410520382e-05
step: 60, loss: 0.0007228404865600169
step: 70, loss: 8.944283035816625e-05
step: 80, loss: 7.697749970247969e-05
step: 90, loss: 5.1568946219049394e-05
step: 100, loss: 6.558569293702021e-05
step: 110, loss: 4.710044231615029e-05
step: 120, loss: 5.559196506510489e-05
step: 130, loss: 7.496495527448133e-05
step: 140, loss: 8.117485413094983e-05
step: 150, loss: 0.00011743311188183725
step: 160, loss: 0.0007654618821106851

```

```

step: 170, loss: 7.69274847698398e-05
step: 180, loss: 9.676686022430658e-05
step: 190, loss: 0.00010521246440475807
step: 200, loss: 0.00019701014389283955
step: 210, loss: 0.00026922280085273087
step: 220, loss: 0.0005513057112693787
step: 230, loss: 0.00034347266773693264
/usr/local/lib/python3.7/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=0.9843749999999999, f1=0.9755011135857461,
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: 0.000272515433607623
step: 10, loss: 0.0005972808576188982
step: 20, loss: 7.516513142036274e-05
step: 30, loss: 0.00015937462740112096
step: 40, loss: 0.000168986851349473
step: 50, loss: 0.00014828269195277244
step: 60, loss: 0.00032138434471562505
step: 70, loss: 7.591494068037719e-05
step: 80, loss: 6.989908433752134e-05
step: 90, loss: 0.0001401997433276847
step: 100, loss: 8.566886390326545e-05
step: 110, loss: 8.930488547775894e-05
step: 120, loss: 0.0029674351681023836
step: 130, loss: 0.0005772713921032846
step: 140, loss: 8.608630014350638e-05
step: 150, loss: 0.019501306116580963
step: 160, loss: 3.706179995788261e-05
step: 170, loss: 0.00012585798685904592
step: 180, loss: 0.00010521014337427914
step: 190, loss: 3.734858182724565e-05
step: 200, loss: 0.0001660722919041291
step: 210, loss: 4.6929089876357466e-05
step: 220, loss: 6.750834290869534e-05
step: 230, loss: 5.846022759214975e-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=0.984304932735426, f1=0.978675645342312,
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: 8.653797704027966e-05

```

```

step: 10, loss: 3.066963836317882e-05
step: 20, loss: 0.00023774182773195207
step: 30, loss: 7.263653969857842e-05
step: 40, loss: 7.067117985570803e-05
step: 50, loss: 0.0001660140696913004
step: 60, loss: 0.00012073584366589785
step: 70, loss: 3.979229586548172e-05
step: 80, loss: 0.00010805492638610303
step: 90, loss: 9.508543735137209e-05
step: 100, loss: 0.0007361210882663727
step: 110, loss: 7.2455448389519e-05
step: 120, loss: 0.00022607902064919472
step: 130, loss: 8.495755173498765e-05
step: 140, loss: 8.48116833367385e-05
step: 150, loss: 6.256200140342116e-05
step: 160, loss: 0.0026001569349318743
step: 170, loss: 4.6972658310551196e-05
step: 180, loss: 0.00011026675201719627
step: 190, loss: 0.0003859246789943427
step: 200, loss: 5.4977113904897124e-05
step: 210, loss: 0.00036259787157177925
step: 220, loss: 4.7017718316055834e-05
step: 230, loss: 3.0457238608505577e-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=0.9855072463768116, f1=0.9810479375696767,
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: 8.31305660540238e-05
step: 10, loss: 8.890658500604331e-05
step: 20, loss: 4.1534731280989945e-05
step: 30, loss: 4.80424496345222e-05
step: 40, loss: 6.505646160803735e-05
step: 50, loss: 0.00011803520465036854
step: 60, loss: 3.4264132409589365e-05
step: 70, loss: 0.00011500095570227131
step: 80, loss: 0.0006780560943298042
step: 90, loss: 3.042753814952448e-05
step: 100, loss: 5.5067375797079876e-05
step: 110, loss: 3.979955363320187e-05
step: 120, loss: 0.00012554449494928122
step: 130, loss: 5.020293974666856e-05
step: 140, loss: 6.373351789079607e-05
step: 150, loss: 8.739723853068426e-05
step: 160, loss: 0.00016258392133750021

```

```

step: 170, loss: 3.574644142645411e-05
step: 180, loss: 4.346451532910578e-05
step: 190, loss: 0.00014705448120366782
step: 200, loss: 0.0004076550540048629
step: 210, loss: 0.00034228444565087557
step: 220, loss: 0.0007937828195281327
step: 230, loss: 6.21062281425111e-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=0.9855072463768116, f1=0.9777777777777777,
best_f1=0.9774266365688488
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 [00:07, 331.00it/s]
load_f1 = 0.9831649831649831
real_f1 = 0.9842696629213483
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:11, 379.91it/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 distilbert \
  --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 distilbert \
  --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: 0.8015668988227844
step: 10, loss: 0.4433898627758026
step: 20, loss: 0.4983620047569275
step: 30, loss: 0.4720902442932129
step: 40, loss: 0.41907060146331787
step: 50, loss: 0.3966452479362488
step: 60, loss: 0.414578914642334
step: 70, loss: 0.27328911423683167
step: 80, loss: 0.23059490323066711
step: 90, loss: 0.11534065753221512
step: 100, loss: 0.31670287251472473
step: 110, loss: 0.1115921214222908
step: 120, loss: 0.058454256504774094
step: 130, loss: 0.04108607769012451
step: 140, loss: 0.24266774952411652
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.023837478831410408
step: 160, loss: 0.1337270587682724
step: 170, loss: 0.42858898639678955
step: 180, loss: 0.11177460104227066
step: 190, loss: 0.04127928614616394
step: 200, loss: 0.24884529411792755

```



```

step: 210, loss: 0.08360590040683746
step: 220, loss: 0.16107484698295593
step: 230, loss: 0.06283603608608246
step: 240, loss: 0.12402033805847168
step: 250, loss: 0.09540601819753647
step: 260, loss: 0.034810103476047516
step: 270, loss: 0.030571334064006805
step: 280, loss: 0.23214773833751678
step: 290, loss: 0.020358087494969368
step: 300, loss: 0.08627106994390488
step: 310, loss: 0.11009883135557175
step: 320, loss: 0.07088520377874374
step: 330, loss: 0.18480217456817627
step: 340, loss: 0.19793912768363953
step: 350, loss: 0.14039725065231323
step: 360, loss: 0.09345652908086777
step: 370, loss: 0.10778676718473434
step: 380, loss: 0.26064828038215637
step: 390, loss: 0.0736495852470398
step: 400, loss: 0.0972292497754097
step: 410, loss: 0.08692849427461624
step: 420, loss: 0.05869833752512932
step: 430, loss: 0.09307004511356354
step: 440, loss: 0.08822275698184967
step: 450, loss: 0.0825580433011055
step: 460, loss: 0.029666010290384293
step: 470, loss: 0.281985342502594
step: 480, loss: 0.23791873455047607
step: 490, loss: 0.050482943654060364
step: 500, loss: 0.04225622117519379
step: 510, loss: 0.08425330370664597
step: 520, loss: 0.14965960383415222
step: 530, loss: 0.1445934772491455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.908093278463649, f1=0.9076853115052297,
best_f1=0.9076853115052297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1635744720697403
step: 10, loss: 0.12234705686569214
step: 20, loss: 0.11662682145833969

```

step: 30, loss: 0.075285404920578
step: 40, loss: 0.024757055565714836
step: 50, loss: 0.12472674995660782
step: 60, loss: 0.19761653244495392
step: 70, loss: 0.09447881579399109
step: 80, loss: 0.014421743340790272
step: 90, loss: 0.03242560848593712
step: 100, loss: 0.29251477122306824
step: 110, loss: 0.10742182284593582
step: 120, loss: 0.12259852141141891
step: 130, loss: 0.05328384041786194
step: 140, loss: 0.07445138692855835
step: 150, loss: 0.07891406863927841
step: 160, loss: 0.022718727588653564
step: 170, loss: 0.11319611221551895
step: 180, loss: 0.003642972791567445
step: 190, loss: 0.05515695735812187
step: 200, loss: 0.038093991577625275
step: 210, loss: 0.041830696165561676
step: 220, loss: 0.18378430604934692
step: 230, loss: 0.03538836911320686
step: 240, loss: 0.11801452934741974
step: 250, loss: 0.10717681050300598
step: 260, loss: 0.04840077459812164
step: 270, loss: 0.2408243864774704
step: 280, loss: 0.1403304487466812
step: 290, loss: 0.14169207215309143
step: 300, loss: 0.037114109843969345
step: 310, loss: 0.07400522381067276
step: 320, loss: 0.10782275348901749
step: 330, loss: 0.060025088489055634
step: 340, loss: 0.023055393248796463
step: 350, loss: 0.04768810793757439
step: 360, loss: 0.08926109969615936
step: 370, loss: 0.00510754156857729
step: 380, loss: 0.09275931864976883
step: 390, loss: 0.03561994433403015
step: 400, loss: 0.16660457849502563
step: 410, loss: 0.03162667900323868
step: 420, loss: 0.05037368834018707
step: 430, loss: 0.030308451503515244
step: 440, loss: 0.030868375673890114
step: 450, loss: 0.023966223001480103
step: 460, loss: 0.16215252876281738
step: 470, loss: 0.11945296823978424
step: 480, loss: 0.23693659901618958
step: 490, loss: 0.07012894004583359
step: 500, loss: 0.03853699192404747

```

step: 510, loss: 0.023570187389850616
step: 520, loss: 0.08262629806995392
step: 530, loss: 0.09197576344013214
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9090909090909091, f1=0.9089201877934272,
best_f1=0.9089201877934272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03364903852343559
step: 10, loss: 0.08238589763641357
step: 20, loss: 0.21271686255931854
step: 30, loss: 0.10437895357608795
step: 40, loss: 0.005388392601162195
step: 50, loss: 0.036851394921541214
step: 60, loss: 0.006796968635171652
step: 70, loss: 0.0680994763970375
step: 80, loss: 0.012880392372608185
step: 90, loss: 0.11752966791391373
step: 100, loss: 0.13181796669960022
step: 110, loss: 0.046784404665231705
step: 120, loss: 0.056908831000328064
step: 130, loss: 0.1353113204240799
step: 140, loss: 0.09745151549577713
step: 150, loss: 0.006236321292817593
step: 160, loss: 0.025169236585497856
step: 170, loss: 0.13258624076843262
step: 180, loss: 0.006473560817539692
step: 190, loss: 0.016147399321198463
step: 200, loss: 0.052590902894735336
step: 210, loss: 0.06976525485515594
step: 220, loss: 0.08554373681545258
step: 230, loss: 0.023543598130345345
step: 240, loss: 0.009976116009056568
step: 250, loss: 0.007683977950364351
step: 260, loss: 0.002425788203254342
step: 270, loss: 0.012104571796953678
step: 280, loss: 0.05008845031261444
step: 290, loss: 0.09866932779550552
step: 300, loss: 0.10180751234292984
step: 310, loss: 0.015613515861332417
step: 320, loss: 0.116200290620327

```

```

step: 330, loss: 0.004953320138156414
step: 340, loss: 0.0126670366153121
step: 350, loss: 0.044261664152145386
step: 360, loss: 0.010835806839168072
step: 370, loss: 0.0011578821577131748
step: 380, loss: 0.0016625396674498916
step: 390, loss: 0.0028778319247066975
step: 400, loss: 0.09612978994846344
step: 410, loss: 0.009796097874641418
step: 420, loss: 0.017177697271108627
step: 430, loss: 0.02359262853860855
step: 440, loss: 0.02519870549440384
step: 450, loss: 0.11169914156198502
step: 460, loss: 0.16564643383026123
step: 470, loss: 0.028068190440535545
step: 480, loss: 0.020471761003136635
step: 490, loss: 0.005665647331625223
step: 500, loss: 0.13440267741680145
step: 510, loss: 0.014700235798954964
step: 520, loss: 0.025805003941059113
step: 530, loss: 0.0389508418738842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9118738404452691, f1=0.9092592592592593,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018091190606355667
step: 10, loss: 0.08520743250846863
step: 20, loss: 0.11023766547441483
step: 30, loss: 0.08039765059947968
step: 40, loss: 0.06253930181264877
step: 50, loss: 0.012348257005214691
step: 60, loss: 0.007327646482735872
step: 70, loss: 0.0017444289987906814
step: 80, loss: 0.02344462275505066
step: 90, loss: 0.021963922306895256
step: 100, loss: 0.003045242978259921
step: 110, loss: 0.0014350912533700466
step: 120, loss: 0.0008063264540396631
step: 130, loss: 0.031115882098674774
step: 140, loss: 0.014394264668226242

```

```

step: 150, loss: 0.0019156085327267647
step: 160, loss: 0.07136973738670349
step: 170, loss: 0.0033848348539322615
step: 180, loss: 0.035294823348522186
step: 190, loss: 0.06509023904800415
step: 200, loss: 0.0008696488221175969
step: 210, loss: 0.08477317541837692
step: 220, loss: 0.04779260233044624
step: 230, loss: 0.330922394990921
step: 240, loss: 0.039267104119062424
step: 250, loss: 0.008882416412234306
step: 260, loss: 0.07587681710720062
step: 270, loss: 0.042910508811473846
step: 280, loss: 0.004983692895621061
step: 290, loss: 0.0025917126331478357
step: 300, loss: 0.05390138551592827
step: 310, loss: 0.006551786791533232
step: 320, loss: 0.015247110277414322
step: 330, loss: 0.20590181648731232
step: 340, loss: 0.036268673837184906
step: 350, loss: 0.12763766944408417
step: 360, loss: 0.01039565447717905
step: 370, loss: 0.015853261575102806
step: 380, loss: 0.016911853104829788
step: 390, loss: 0.048386264592409134
step: 400, loss: 0.03575609624385834
step: 410, loss: 0.07732705026865005
step: 420, loss: 0.002775361994281411
step: 430, loss: 0.003734740661457181
step: 440, loss: 0.02810184843838215
step: 450, loss: 0.02527523785829544
step: 460, loss: 0.02556423470377922
step: 470, loss: 0.12038984894752502
step: 480, loss: 0.012694505974650383
step: 490, loss: 0.02231069840490818
step: 500, loss: 0.030850540846586227
step: 510, loss: 0.009699856862425804
step: 520, loss: 0.02142782136797905
step: 530, loss: 0.0007568910950794816
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.9086778736937755, f1=0.9059518400726941,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.0022740340791642666
step: 10, loss: 0.0023119242396205664
step: 20, loss: 0.006150748580694199
step: 30, loss: 0.004954617470502853
step: 40, loss: 0.0022172704339027405
step: 50, loss: 0.002055257558822632
step: 60, loss: 0.0010016094893217087
step: 70, loss: 0.012715407647192478
step: 80, loss: 0.0011862224200740457
step: 90, loss: 0.05786813795566559
step: 100, loss: 0.0009901025332510471
step: 110, loss: 0.0060372790321707726
step: 120, loss: 0.009769415482878685
step: 130, loss: 0.0024843697901815176
step: 140, loss: 0.0005949846818111837
step: 150, loss: 0.0007752066012471914
step: 160, loss: 0.14516018331050873
step: 170, loss: 0.016924334689974785
step: 180, loss: 0.0010661690030246973
step: 190, loss: 0.0010927689727395773
step: 200, loss: 0.0012571201659739017
step: 210, loss: 0.004496603272855282
step: 220, loss: 0.00035123236011713743
step: 230, loss: 0.05558747425675392
step: 240, loss: 0.02377576008439064
step: 250, loss: 0.00602414645254612
step: 260, loss: 0.005786759313195944
step: 270, loss: 0.009490675292909145
step: 280, loss: 0.00465143658220768
step: 290, loss: 0.001005747471936047
step: 300, loss: 0.14752471446990967
step: 310, loss: 0.0006414172239601612
step: 320, loss: 0.005426892079412937
step: 330, loss: 0.012850082479417324
step: 340, loss: 0.002037316095083952
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 350, loss: 0.07081396877765656
step: 360, loss: 0.024063771590590477
step: 370, loss: 0.03938323259353638
step: 380, loss: 0.07842055708169937
step: 390, loss: 0.001071310369297862
step: 400, loss: 0.001127993338741362
step: 410, loss: 0.0006920213345438242
step: 420, loss: 0.011853414587676525
step: 430, loss: 0.01428088080137968
step: 440, loss: 0.0009037298150360584
step: 450, loss: 0.0009221715154126287
step: 460, loss: 0.0006539793103002012

```
step: 470, loss: 0.0006431074580177665
step: 480, loss: 0.001093561644665897
step: 490, loss: 0.004906331654638052
step: 500, loss: 0.09887981414794922
step: 510, loss: 0.006278504151850939
step: 520, loss: 0.0008272022241726518
step: 530, loss: 0.099171943962574
/usr/local/lib/python3.7/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=0.9052242256125751, f1=0.9154411764705882,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015744702890515327
step: 10, loss: 0.005128773860633373
step: 20, loss: 0.0013393588596954942
step: 30, loss: 0.031061459332704544
step: 40, loss: 0.003123676171526313
step: 50, loss: 0.0009767950978130102
step: 60, loss: 0.0013387809740379453
step: 70, loss: 0.003406330244615674
step: 80, loss: 0.004735324531793594
step: 90, loss: 0.01605076901614666
step: 100, loss: 0.004119398538023233
step: 110, loss: 0.001927731093019247
step: 120, loss: 0.0013788630021736026
step: 130, loss: 0.0002141161821782589
step: 140, loss: 0.005159108899533749
step: 150, loss: 0.002387887565419078
step: 160, loss: 0.0002612634270917624
step: 170, loss: 0.0001816911535570398
step: 180, loss: 0.00013849910465069115
step: 190, loss: 0.019726144149899483
step: 200, loss: 0.00031153467716649175
step: 210, loss: 0.000594346085563302
step: 220, loss: 0.009691428393125534
step: 230, loss: 0.0005129874334670603
step: 240, loss: 0.0015522395260632038
step: 250, loss: 0.00033944917959161103
step: 260, loss: 0.0011818427592515945
step: 270, loss: 0.0009655195171944797
step: 280, loss: 0.12353097647428513
step: 290, loss: 0.000869478564709425
step: 300, loss: 0.0029003205709159374
step: 310, loss: 0.0019707223400473595
step: 320, loss: 0.0006749979802407324
```

```

step: 330, loss: 0.00433393893763423
step: 340, loss: 0.0023760180920362473
step: 350, loss: 0.10230743139982224
step: 360, loss: 0.0006696871132589877
step: 370, loss: 0.004264051094651222
step: 380, loss: 0.0007566502317786217
step: 390, loss: 0.09616102278232574
step: 400, loss: 0.0005824083928018808
step: 410, loss: 0.0007598886149935424
step: 420, loss: 0.14506705105304718
step: 430, loss: 0.002115207491442561
step: 440, loss: 0.002268922282382846
step: 450, loss: 0.0008415071642957628
step: 460, loss: 0.0011463541304692626
step: 470, loss: 0.020504148676991463
step: 480, loss: 0.06834492087364197
step: 490, loss: 0.0037460578605532646
step: 500, loss: 0.0012403170112520456
step: 510, loss: 0.005517726764082909
step: 520, loss: 0.0018800412071868777
step: 530, loss: 0.1500268280506134
/usr/local/lib/python3.7/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=0.9054863992623329, f1=0.9067405355493998,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007738708518445492
step: 10, loss: 0.004011825192719698
step: 20, loss: 0.001286576152779162
step: 30, loss: 0.0025652749463915825
step: 40, loss: 0.00015151509433053434
step: 50, loss: 0.0012461114674806595
step: 60, loss: 0.11888080090284348
step: 70, loss: 0.0010462702484801412
step: 80, loss: 0.008584434166550636
step: 90, loss: 0.0009830828057602048
step: 100, loss: 0.0013895768206566572
step: 110, loss: 0.004194274544715881
step: 120, loss: 0.0009836637182161212
step: 130, loss: 0.000844435126055032
step: 140, loss: 0.0008885706192813814
step: 150, loss: 0.0001752735988702625
step: 160, loss: 0.0016201066318899393
step: 170, loss: 0.17422114312648773
step: 180, loss: 0.00020035436318721622

```



```

step: 190, loss: 0.0005486299050971866
step: 200, loss: 0.0074887811206281185
step: 210, loss: 0.00466389674693346
step: 220, loss: 0.00045526016037911177
step: 230, loss: 0.06588653475046158
step: 240, loss: 0.0022928931284695864
step: 250, loss: 0.0005568856722675264
step: 260, loss: 0.05773273855447769
step: 270, loss: 0.005284425336867571
step: 280, loss: 0.016517652198672295
step: 290, loss: 0.058898016810417175
step: 300, loss: 0.0008065786096267402
step: 310, loss: 0.0010187345324084163
step: 320, loss: 0.011286181397736073
step: 330, loss: 0.00014086243754718453
step: 340, loss: 0.1377863585948944
step: 350, loss: 0.004278677515685558
step: 360, loss: 0.0017649347428232431
step: 370, loss: 0.0035000841598957777
step: 380, loss: 0.0006861714064143598
step: 390, loss: 0.001707727089524269
step: 400, loss: 0.00013078676420263946
step: 410, loss: 0.0007883553043939173
step: 420, loss: 0.006561249028891325
step: 430, loss: 0.01559208519756794
step: 440, loss: 0.0029168729670345783
step: 450, loss: 0.002002985682338476
step: 460, loss: 0.004098803270608187
step: 470, loss: 0.07841745018959045
step: 480, loss: 0.044532518833875656
step: 490, loss: 0.0005062154377810657
step: 500, loss: 0.0023882868699729443
step: 510, loss: 0.0031846945639699697
step: 520, loss: 0.0008296953747048974
step: 530, loss: 0.00031200170633383095
/usr/local/lib/python3.7/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=0.909433962264151, f1=0.8994307400379506,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00047728922800160944
step: 10, loss: 0.0006812381907366216
step: 20, loss: 0.0005548332701437175
step: 30, loss: 0.0002960096171591431
step: 40, loss: 0.0007275805692188442

```

step: 50, loss: 0.0019904784858226776
step: 60, loss: 0.0006368912872858346
step: 70, loss: 0.0004819017485715449
step: 80, loss: 0.0001606924633961171
step: 90, loss: 0.0007188365561887622
step: 100, loss: 0.03994770348072052
step: 110, loss: 0.011977311223745346
step: 120, loss: 0.0007050312706269324
step: 130, loss: 0.0006118597229942679
step: 140, loss: 5.955228698439896e-05
step: 150, loss: 0.001989064272493124
step: 160, loss: 0.00033726077526807785
step: 170, loss: 0.00024542261962778866
step: 180, loss: 0.0007180693792179227
step: 190, loss: 0.0001294064859393984
step: 200, loss: 0.0003283853002358228
step: 210, loss: 0.00018594168068375438
step: 220, loss: 0.007621414493769407
step: 230, loss: 0.0001705247996142134
step: 240, loss: 0.00010461585043231025
step: 250, loss: 0.018777815625071526
step: 260, loss: 0.03646114841103554
step: 270, loss: 0.0001501388178439811
step: 280, loss: 0.02699115313589573
step: 290, loss: 0.003548906184732914
step: 300, loss: 9.62605481618084e-05
step: 310, loss: 0.001217918354086578
step: 320, loss: 0.0005452413461171091
step: 330, loss: 0.002250111661851406
step: 340, loss: 0.0004431677225511521
step: 350, loss: 0.0010341573506593704
step: 360, loss: 0.0049090757966041565
step: 370, loss: 0.0009095208370126784
step: 380, loss: 0.0003228996938560158
step: 390, loss: 0.00016027603123802692
step: 400, loss: 0.10015103220939636
step: 410, loss: 0.00017141785065177828
step: 420, loss: 0.0009933356195688248
step: 430, loss: 0.00019757177506107837
step: 440, loss: 0.0005154658574610949
step: 450, loss: 0.0008874867344275117
step: 460, loss: 0.013053635135293007
step: 470, loss: 0.0022448855452239513
step: 480, loss: 0.015372530557215214
step: 490, loss: 0.0018318088259547949
step: 500, loss: 0.01760317198932171
step: 510, loss: 9.110527025768533e-05
step: 520, loss: 0.0002697074378374964

```
step: 530, loss: 0.0005447129951789975
/usr/local/lib/python3.7/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=0.8986064392119173, f1=0.8911465892597968,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016371645324397832
step: 10, loss: 0.0001273148664040491
step: 20, loss: 0.0010258350521326065
step: 30, loss: 0.01547228917479515
step: 40, loss: 0.07138422876596451
step: 50, loss: 0.00895835179835558
step: 60, loss: 0.0010151737369596958
step: 70, loss: 0.023025384172797203
step: 80, loss: 0.00015414973313454539
step: 90, loss: 0.0028678737580776215
step: 100, loss: 0.0018300500232726336
step: 110, loss: 0.00029032593010924757
step: 120, loss: 0.00016695164958946407
step: 130, loss: 0.006400061771273613
step: 140, loss: 0.006201942916959524
step: 150, loss: 0.018840571865439415
step: 160, loss: 0.00018564755737315863
step: 170, loss: 0.0003024559991899878
step: 180, loss: 0.0005003103287890553
step: 190, loss: 0.002008007373660803
step: 200, loss: 0.0005185359623283148
step: 210, loss: 8.099174010567367e-05
step: 220, loss: 5.026230064686388e-05
step: 230, loss: 0.00010799484880408272
step: 240, loss: 0.00015719352813903242
step: 250, loss: 0.00023231824161484838
step: 260, loss: 0.0027767117135226727
step: 270, loss: 0.0025311000645160675
step: 280, loss: 0.00011376726615708321
step: 290, loss: 5.4598138376604766e-05
step: 300, loss: 0.001585114048793912
step: 310, loss: 0.0002016408834606409
step: 320, loss: 0.00025155249750241637
step: 330, loss: 0.00035378008033148944
step: 340, loss: 0.0001841826451709494
step: 350, loss: 0.00010804258636198938
step: 360, loss: 0.00044982507824897766
step: 370, loss: 0.0003827604232355952
step: 380, loss: 4.6868808567523956e-05
```

```

step: 390, loss: 6.465529440902174e-05
step: 400, loss: 0.007334086578339338
step: 410, loss: 0.06828516721725464
step: 420, loss: 5.924570723436773e-05
step: 430, loss: 5.074220825918019e-05
step: 440, loss: 0.00016168384172488004
step: 450, loss: 5.27763731952291e-05
step: 460, loss: 0.0012535977875813842
step: 470, loss: 0.001164085348136723
step: 480, loss: 0.0012362132547423244
step: 490, loss: 0.002170414896681905
step: 500, loss: 0.003760626772418618
step: 510, loss: 0.0005867494037374854
step: 520, loss: 0.0005639904993586242
step: 530, loss: 0.00600047130137682
/usr/local/lib/python3.7/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=0.908673703533731, f1=0.9167429094236048,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014150640927255154
step: 10, loss: 3.975005529355258e-05
step: 20, loss: 0.00017797584587242454
step: 30, loss: 8.240703755291179e-05
step: 40, loss: 0.0005437403451651335
step: 50, loss: 0.0006688365247100592
step: 60, loss: 9.696525376057252e-05
step: 70, loss: 0.014114673249423504
step: 80, loss: 0.0009445353643968701
step: 90, loss: 0.001663287985138595
step: 100, loss: 0.0369722954928875
step: 110, loss: 0.004148394335061312
step: 120, loss: 0.00044312674435786903
step: 130, loss: 6.451496301451698e-05
step: 140, loss: 0.00033808592706918716
step: 150, loss: 0.012701950035989285
step: 160, loss: 0.0002205323107773438
step: 170, loss: 4.34120993304532e-05
step: 180, loss: 0.005561197642236948
step: 190, loss: 0.00015528377844020724
step: 200, loss: 0.004979541525244713
step: 210, loss: 0.0001512483722763136
step: 220, loss: 4.301730587030761e-05
step: 230, loss: 5.096535460324958e-05
step: 240, loss: 0.00010835708235390484

```

```

step: 250, loss: 0.00308537227101624
step: 260, loss: 0.0052301036193966866
step: 270, loss: 0.00015611819981131703
step: 280, loss: 9.341367695014924e-05
step: 290, loss: 5.783591768704355e-05
step: 300, loss: 0.003658928209915757
step: 310, loss: 0.00017420813674107194
step: 320, loss: 2.920163024100475e-05
step: 330, loss: 5.603733370662667e-05
step: 340, loss: 0.0011857408098876476
step: 350, loss: 7.729097706032917e-05
step: 360, loss: 4.90646343678236e-05
step: 370, loss: 0.0008384347311221063
step: 380, loss: 4.595557402353734e-05
step: 390, loss: 7.895704766269773e-05
step: 400, loss: 5.662070179823786e-05
step: 410, loss: 0.07801194489002228
step: 420, loss: 0.0030646363738924265
step: 430, loss: 6.917420250829309e-05
step: 440, loss: 4.6157547330949455e-05
step: 450, loss: 0.001839947421103716
step: 460, loss: 0.0021916311234235764
step: 470, loss: 9.971218241844326e-05
step: 480, loss: 0.0001566661085234955
step: 490, loss: 0.017067670822143555
step: 500, loss: 0.0008613317622803152
step: 510, loss: 8.840232476359233e-05
step: 520, loss: 0.00014096777886152267
step: 530, loss: 3.619323979364708e-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=0.9100138440239963, f1=0.9134438305709024,
best_f1=0.9092592592592593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000288584444206208
step: 10, loss: 0.000342927232850343
step: 20, loss: 0.00010493597073946148
step: 30, loss: 0.007659786380827427
step: 40, loss: 0.00016665791918057948
step: 50, loss: 0.0003264600527472794
step: 60, loss: 4.5333628804655746e-05
step: 70, loss: 3.532560731400736e-05
step: 80, loss: 0.0007927102269604802
step: 90, loss: 4.531586091616191e-05
step: 100, loss: 0.000139359719469212

```

```
step: 110, loss: 6.160556949907914e-05
step: 120, loss: 0.0009541409090161324
step: 130, loss: 9.15236450964585e-05
step: 140, loss: 3.0460325433523394e-05
step: 150, loss: 4.386510408949107e-05
step: 160, loss: 6.349917384795845e-05
step: 170, loss: 0.00011558182450244203
step: 180, loss: 4.7817982704145834e-05
step: 190, loss: 7.63397547416389e-05
step: 200, loss: 0.0013419484021142125
step: 210, loss: 0.0002511379134375602
step: 220, loss: 7.941363583086058e-05
step: 230, loss: 3.8938822399359196e-05
step: 240, loss: 5.17101252626162e-05
step: 250, loss: 0.000851538497954607
step: 260, loss: 3.406237010494806e-05
step: 270, loss: 0.000185455268365331
step: 280, loss: 0.00045390697778202593
step: 290, loss: 9.883505845209584e-05
step: 300, loss: 5.957367829978466e-05
step: 310, loss: 0.0004979007644578815
step: 320, loss: 0.0004084636748302728
step: 330, loss: 0.0001374171260977164
step: 340, loss: 0.029756445437669754
step: 350, loss: 0.014180128462612629
step: 360, loss: 0.015456081368029118
step: 370, loss: 3.155498416163027e-05
step: 380, loss: 2.0749472241732292e-05
step: 390, loss: 0.00021236717293504626
step: 400, loss: 2.2448091840487905e-05
step: 410, loss: 5.781118306913413e-05
step: 420, loss: 0.0001137325307354331
step: 430, loss: 8.85897025000304e-05
step: 440, loss: 4.980422818334773e-05
step: 450, loss: 3.0985233024694026e-05
step: 460, loss: 0.0011922362027689815
step: 470, loss: 7.645436562597752e-05
step: 480, loss: 4.4279258872848004e-05
step: 490, loss: 0.0008531258208677173
step: 500, loss: 0.03272876888513565
step: 510, loss: 3.484483386273496e-05
step: 520, loss: 3.154476507916115e-05
step: 530, loss: 3.141806155326776e-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=0.9105022831050228, f1=0.9118318867062585,
best_f1=0.9092592592592593
```

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

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

step: 0, loss: 8.76597041497007e-05
step: 10, loss: 3.641721559688449e-05
step: 20, loss: 2.1960115191177465e-05
step: 30, loss: 7.376878784270957e-05
step: 40, loss: 3.468379145488143e-05
step: 50, loss: 3.4322787541896105e-05
step: 60, loss: 4.216817978885956e-05
step: 70, loss: 6.887106428621337e-05
step: 80, loss: 0.00042217384907417
step: 90, loss: 0.0014711320400238037
step: 100, loss: 0.0134600680321455
step: 110, loss: 0.00010357661813031882
step: 120, loss: 0.0014265357749536633
step: 130, loss: 0.00019978625641670078
step: 140, loss: 5.422012327471748e-05
step: 150, loss: 4.7380690375575796e-05
step: 160, loss: 7.467050454579294e-05
step: 170, loss: 3.6226647353032604e-05
step: 180, loss: 0.00013996286725159734
step: 190, loss: 0.00012570539547596127
step: 200, loss: 0.00012152235285611823
step: 210, loss: 0.000766789773479104
step: 220, loss: 0.0002932303468696773
step: 230, loss: 7.461149652954191e-05
step: 240, loss: 8.983931184047833e-05
step: 250, loss: 0.0012917546555399895
step: 260, loss: 0.0030441302806138992
step: 270, loss: 0.00021861046843696386
step: 280, loss: 0.0014686761423945427
step: 290, loss: 8.062426786636934e-05
step: 300, loss: 0.0003799512633122504
step: 310, loss: 0.0001167498849099502
step: 320, loss: 4.489383354666643e-05
step: 330, loss: 2.8702586860163137e-05
step: 340, loss: 2.873596895369701e-05
step: 350, loss: 0.00013887148816138506
step: 360, loss: 0.0012508518993854523
step: 370, loss: 7.935579924378544e-05
step: 380, loss: 0.0003822458675131202
step: 390, loss: 0.00028418793226592243
step: 400, loss: 0.003016019007191062
step: 410, loss: 0.0025421546306461096
step: 420, loss: 6.709455192321911e-05
step: 430, loss: 0.00634046969935298
step: 440, loss: 4.507108315010555e-05

```

step: 450, loss: 0.00019058525504078716
step: 460, loss: 0.0003583567449823022
step: 470, loss: 2.723463148868177e-05
step: 480, loss: 4.819735841010697e-05
step: 490, loss: 2.4057417249423452e-05
step: 500, loss: 0.00016051478451117873
step: 510, loss: 9.321246761828661e-05
step: 520, loss: 0.0002719072508625686
step: 530, loss: 6.789210601709783e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 12: dev_f1=0.9124594719777674, f1=0.9130234698573401,
best_f1=0.9130234698573401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012802870478481054
step: 10, loss: 0.0006044319598004222
step: 20, loss: 3.396538886590861e-05
step: 30, loss: 7.9010846093297e-05
step: 40, loss: 2.826665513566695e-05
step: 50, loss: 0.0002197989379055798
step: 60, loss: 4.764917684951797e-05
step: 70, loss: 5.7269939134130254e-05
step: 80, loss: 0.00012343018897809088
step: 90, loss: 5.727107418351807e-05
step: 100, loss: 2.9778773750877008e-05
step: 110, loss: 0.0005849956069141626
step: 120, loss: 7.876814197516069e-05
step: 130, loss: 0.00013856875011697412
step: 140, loss: 5.056336522102356e-05
step: 150, loss: 9.771491750143468e-05
step: 160, loss: 4.197258022031747e-05
step: 170, loss: 4.6494740672642365e-05
step: 180, loss: 2.98795730486745e-05
step: 190, loss: 0.00014493605704046786
step: 200, loss: 3.475874109426513e-05
step: 210, loss: 4.2350082367192954e-05
step: 220, loss: 3.8226662582019344e-05
step: 230, loss: 5.669979509548284e-05
step: 240, loss: 4.362436811788939e-05
step: 250, loss: 0.002004502108320594
step: 260, loss: 2.36215037148213e-05

```



```

step: 270, loss: 2.6136094675166532e-05
step: 280, loss: 6.135896546766162e-05
step: 290, loss: 4.209410690236837e-05
step: 300, loss: 4.2019055399578065e-05
step: 310, loss: 0.0032915500923991203
step: 320, loss: 6.608395779039711e-05
step: 330, loss: 8.321487985085696e-05
step: 340, loss: 4.279181666788645e-05
step: 350, loss: 0.003570662811398506
step: 360, loss: 0.002055651042610407
step: 370, loss: 2.8038808522978798e-05
step: 380, loss: 0.06495507806539536
step: 390, loss: 2.8833004762418568e-05
step: 400, loss: 1.9456892914604396e-05
step: 410, loss: 8.664381311973557e-05
step: 420, loss: 2.4072280211839825e-05
step: 430, loss: 0.00017824476526584476
step: 440, loss: 8.85301997186616e-05
step: 450, loss: 3.62453174602706e-05
step: 460, loss: 2.0712210243800655e-05
step: 470, loss: 3.503088737488724e-05
step: 480, loss: 2.6355481168138795e-05
step: 490, loss: 5.883228004677221e-05
step: 500, loss: 0.00012167952809249982
step: 510, loss: 2.567037518019788e-05
step: 520, loss: 4.0400751458946615e-05
step: 530, loss: 0.00018574329442344606
/usr/local/lib/python3.7/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=0.9109461966604824, f1=0.9190189726978251,
best_f1=0.9130234698573401
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001632978382986039
step: 10, loss: 4.710893699666485e-05
step: 20, loss: 2.4504392058588564e-05
step: 30, loss: 9.592499554855749e-05
step: 40, loss: 2.236987347714603e-05
step: 50, loss: 6.752838817192242e-05
step: 60, loss: 4.2201849282719195e-05
step: 70, loss: 3.97696130676195e-05
step: 80, loss: 3.724971611518413e-05
step: 90, loss: 1.8939048459287733e-05
step: 100, loss: 7.139497029129416e-05
step: 110, loss: 2.5282635760959238e-05
step: 120, loss: 7.155705679906532e-05

```

```

step: 130, loss: 2.3777809474267997e-05
step: 140, loss: 0.00028545319219119847
step: 150, loss: 4.31591433880385e-05
step: 160, loss: 1.9836747014778666e-05
step: 170, loss: 0.0008755846065469086
step: 180, loss: 0.00014449574518948793
step: 190, loss: 8.530059858458117e-05
step: 200, loss: 4.7988203732529655e-05
step: 210, loss: 3.8822927308501676e-05
step: 220, loss: 0.00022921945492271334
step: 230, loss: 4.4329168304102495e-05
step: 240, loss: 6.580037734238431e-05
step: 250, loss: 4.284912574803457e-05
step: 260, loss: 1.56385558511829e-05
step: 270, loss: 8.452207111986354e-05
step: 280, loss: 1.9166289348504506e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 290, loss: 0.05517970770597458
step: 300, loss: 2.24067425733665e-05
step: 310, loss: 3.549434040905908e-05
step: 320, loss: 3.4706714359344915e-05
step: 330, loss: 2.3598719053552486e-05
step: 340, loss: 3.847313564619981e-05
step: 350, loss: 2.5647586880950257e-05
step: 360, loss: 3.0653605790575966e-05
step: 370, loss: 2.2932374122319743e-05
step: 380, loss: 1.6633179257041775e-05
step: 390, loss: 0.00031554230372421443
step: 400, loss: 3.1839073926676065e-05
step: 410, loss: 3.514786294545047e-05
step: 420, loss: 5.586723273154348e-05
step: 430, loss: 0.0009890178916975856
step: 440, loss: 1.7407965060556307e-05
step: 450, loss: 2.2693979190080427e-05
step: 460, loss: 2.8083995857741684e-05
step: 470, loss: 2.4086662961053662e-05
step: 480, loss: 3.18863385473378e-05
step: 490, loss: 3.777512756641954e-05
step: 500, loss: 2.6459603759576567e-05
step: 510, loss: 0.0001485592219978571
step: 520, loss: 3.600718264351599e-05
step: 530, loss: 3.48740431945771e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```
warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.9127826488232579, f1=0.9183955739972337,
best_f1=0.9183955739972337
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8284448490012437e-05
step: 10, loss: 0.022492388263344765
step: 20, loss: 2.4508162823622115e-05
step: 30, loss: 0.0004943721578456461
step: 40, loss: 1.8160320905735716e-05
step: 50, loss: 3.1073446734808385e-05
step: 60, loss: 2.7616406441666186e-05
step: 70, loss: 3.2884065149119124e-05
step: 80, loss: 2.848563963198103e-05
step: 90, loss: 3.2887292036321014e-05
step: 100, loss: 2.149071588064544e-05
step: 110, loss: 0.0036646926309913397
step: 120, loss: 2.9976294172229245e-05
step: 130, loss: 2.5664796339697205e-05
step: 140, loss: 2.0835070245084353e-05
step: 150, loss: 0.00018887165060732514
step: 160, loss: 2.8821563319070265e-05
step: 170, loss: 4.7770958190085366e-05
step: 180, loss: 0.0015119147719815373
step: 190, loss: 3.385071613593027e-05
step: 200, loss: 4.540775262285024e-05
step: 210, loss: 6.325206049950793e-05
step: 220, loss: 0.0001307096827076748
step: 230, loss: 3.3085147151723504e-05
step: 240, loss: 3.285171260358766e-05
step: 250, loss: 6.84559709043242e-05
step: 260, loss: 1.8391443518339656e-05
step: 270, loss: 4.9020160076906905e-05
step: 280, loss: 4.691938011092134e-05
step: 290, loss: 4.070829163538292e-05
step: 300, loss: 0.0011673050466924906
step: 310, loss: 0.00020174923702143133
step: 320, loss: 5.4407490097219124e-05
step: 330, loss: 4.325558984419331e-05
step: 340, loss: 0.0017853562021628022
step: 350, loss: 6.337727973004803e-05
step: 360, loss: 5.150356810190715e-05
step: 370, loss: 2.185181983804796e-05
step: 380, loss: 2.1680372810806148e-05
step: 390, loss: 2.4604272766737267e-05
step: 400, loss: 3.049334372917656e-05
step: 410, loss: 5.0327576900599524e-05
```

```

step: 420, loss: 2.7361511456547305e-05
step: 430, loss: 1.931529368448537e-05
step: 440, loss: 4.7757137508597225e-05
step: 450, loss: 7.591745816171169e-05
step: 460, loss: 1.9639295715023763e-05
step: 470, loss: 2.2045640434953384e-05
step: 480, loss: 1.6316547771566547e-05
step: 490, loss: 4.293245729058981e-05
step: 500, loss: 4.8616107960697263e-05
step: 510, loss: 3.6255580198485404e-05
step: 520, loss: 2.9290165912243538e-05
step: 530, loss: 2.715970913413912e-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=0.9118058796080262, f1=0.9136724218385441,
best_f1=0.9183955739972337
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 [00:14, 408.38it/s]

```

```
load_f1 = 0.908921933085502
real_f1 = 0.9075630252100841
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 [00:10, 422.08it/s]
```

3.1.4 Amazon-Google - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Amazon-Google \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm distilbert \
  --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 distilbert \
  --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: 0.836605429649353
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.06922208517789841
step: 20, loss: 0.3874260187149048
step: 30, loss: 0.3702666759490967
step: 40, loss: 0.517703115940094
step: 50, loss: 0.3140665590763092
step: 60, loss: 0.37785178422927856
step: 70, loss: 0.24371877312660217
step: 80, loss: 0.2639542818069458
step: 90, loss: 0.45573946833610535
step: 100, loss: 0.19021429121494293
step: 110, loss: 0.30478209257125854
step: 120, loss: 0.25250890851020813
step: 130, loss: 0.2439689338207245
step: 140, loss: 0.251078724861145
step: 150, loss: 0.2669485807418823

```

```

step: 160, loss: 0.24314165115356445
step: 170, loss: 0.18437951803207397
step: 180, loss: 0.19263799488544464
step: 190, loss: 0.21223627030849457
step: 200, loss: 0.18077705800533295
step: 210, loss: 0.4480547606945038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5183673469387755, f1=0.5417515274949083,
best_f1=0.5417515274949083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09668213874101639
step: 10, loss: 0.08878421783447266
step: 20, loss: 0.2818831503391266
step: 30, loss: 0.23243197798728943
step: 40, loss: 0.0882980227470398
step: 50, loss: 0.19631646573543549
step: 60, loss: 0.07719609886407852
step: 70, loss: 0.262552410364151
step: 80, loss: 0.23180805146694183
step: 90, loss: 0.174530029296875
step: 100, loss: 0.13208618760108948
step: 110, loss: 0.08335801959037781
step: 120, loss: 0.17245402932167053
step: 130, loss: 0.24174508452415466
step: 140, loss: 0.21694514155387878
step: 150, loss: 0.35282284021377563
step: 160, loss: 0.1550905853509903
step: 170, loss: 0.20202168822288513
step: 180, loss: 0.20504267513751984
step: 190, loss: 0.15579260885715485
step: 200, loss: 0.13886894285678864
step: 210, loss: 0.35639938712120056
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5487804878048781, f1=0.5596707818930041,

```

```

best_f1=0.5596707818930041
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15719519555568695
step: 10, loss: 0.26975366473197937
step: 20, loss: 0.2352542132139206
step: 30, loss: 0.11113858222961426
step: 40, loss: 0.13367024064064026
step: 50, loss: 0.21392972767353058
step: 60, loss: 0.17192286252975464
step: 70, loss: 0.13072550296783447
step: 80, loss: 0.14822271466255188
step: 90, loss: 0.12195441871881485
step: 100, loss: 0.13973474502563477
step: 110, loss: 0.12084471434354782
step: 120, loss: 0.10060571134090424
step: 130, loss: 0.12633845210075378
step: 140, loss: 0.1685561239719391
step: 150, loss: 0.26101985573768616
step: 160, loss: 0.09219039231538773
step: 170, loss: 0.16308283805847168
step: 180, loss: 0.15178725123405457
step: 190, loss: 0.24261808395385742
step: 200, loss: 0.12261800467967987
step: 210, loss: 0.21791350841522217
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5901639344262296, f1=0.5679012345679013,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19438569247722626
step: 10, loss: 0.11303205788135529
step: 20, loss: 0.08867001533508301
step: 30, loss: 0.05313879996538162
step: 40, loss: 0.061870671808719635
step: 50, loss: 0.11542776226997375
step: 60, loss: 0.03383802995085716
step: 70, loss: 0.15962840616703033
step: 80, loss: 0.1566307246685028
step: 90, loss: 0.04263700917363167

```



```

step: 100, loss: 0.09211613237857819
step: 110, loss: 0.1439387947320938
step: 120, loss: 0.07595734298229218
step: 130, loss: 0.10194196552038193
step: 140, loss: 0.30979156494140625
step: 150, loss: 0.2566286325454712
step: 160, loss: 0.03251461684703827
step: 170, loss: 0.03073284775018692
step: 180, loss: 0.24797506630420685
step: 190, loss: 0.16197554767131805
step: 200, loss: 0.07947154343128204
step: 210, loss: 0.02421874925494194
/usr/local/lib/python3.7/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=0.5677966101694915, f1=0.5033707865168541,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04254663735628128
step: 10, loss: 0.049460120499134064
step: 20, loss: 0.03219735622406006
step: 30, loss: 0.10391955077648163
step: 40, loss: 0.2313525676727295
step: 50, loss: 0.18518321216106415
step: 60, loss: 0.08501609414815903
step: 70, loss: 0.09738428145647049
step: 80, loss: 0.02715439349412918
step: 90, loss: 0.01659199967980385
step: 100, loss: 0.10151905566453934
step: 110, loss: 0.010933262296020985
step: 120, loss: 0.06354466825723648
step: 130, loss: 0.19727298617362976
step: 140, loss: 0.06176042929291725
step: 150, loss: 0.07941463589668274
step: 160, loss: 0.040221329778432846
step: 170, loss: 0.03743274137377739
step: 180, loss: 0.08488289266824722
step: 190, loss: 0.04028576612472534
step: 200, loss: 0.17839670181274414
step: 210, loss: 0.06440810114145279
/usr/local/lib/python3.7/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=0.5469061876247505, f1=0.5508474576271185,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02524157240986824
step: 10, loss: 0.030366091057658195
step: 20, loss: 0.019885050132870674
step: 30, loss: 0.22704271972179413
step: 40, loss: 0.05263625457882881
step: 50, loss: 0.037572044879198074
step: 60, loss: 0.05183097720146179
step: 70, loss: 0.006966148968786001
step: 80, loss: 0.12147849798202515
step: 90, loss: 0.1233554556965828
step: 100, loss: 0.007806209847331047
step: 110, loss: 0.02125062793493271
step: 120, loss: 0.07341016829013824
step: 130, loss: 0.03592449426651001
step: 140, loss: 0.08023754507303238
step: 150, loss: 0.01828954927623272
step: 160, loss: 0.08930371701717377
step: 170, loss: 0.029784085229039192
step: 180, loss: 0.005964319221675396
step: 190, loss: 0.011854474432766438
step: 200, loss: 0.007585427258163691
step: 210, loss: 0.0028623647522181273
/usr/local/lib/python3.7/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=0.548856548856549, f1=0.5258620689655173,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029254650697112083
step: 10, loss: 0.01252291351556778
step: 20, loss: 0.006303764879703522
step: 30, loss: 0.040940314531326294
step: 40, loss: 0.051667459309101105
step: 50, loss: 0.0523749515414238
step: 60, loss: 0.09217405319213867
step: 70, loss: 0.019228041172027588
step: 80, loss: 0.0036338388454169035
step: 90, loss: 0.007170497439801693
step: 100, loss: 0.0007253098301589489
step: 110, loss: 0.06578629463911057
step: 120, loss: 0.04733341559767723
step: 130, loss: 0.0022581357043236494
step: 140, loss: 0.010470395907759666
step: 150, loss: 0.02633313462138176

```

```

step: 160, loss: 0.1701655089855194
step: 170, loss: 0.008443654514849186
step: 180, loss: 0.0025142403319478035
step: 190, loss: 0.0119570791721344
step: 200, loss: 0.0068633961491286755
step: 210, loss: 0.04597422853112221
/usr/local/lib/python3.7/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=0.5547785547785548, f1=0.48780487804878053,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05752791836857796
step: 10, loss: 0.013927408494055271
step: 20, loss: 0.059220343828201294
step: 30, loss: 0.02320295013487339
step: 40, loss: 0.05634685233235359
step: 50, loss: 0.002772262552753091
step: 60, loss: 0.06726213544607162
step: 70, loss: 0.0032905186526477337
step: 80, loss: 0.0009611907880753279
step: 90, loss: 0.14173369109630585
step: 100, loss: 0.007419493980705738
step: 110, loss: 0.001852361485362053
step: 120, loss: 0.0026968298479914665
step: 130, loss: 0.006144918035715818
step: 140, loss: 0.0020890706218779087
step: 150, loss: 0.00567190907895565
step: 160, loss: 0.07172340899705887
step: 170, loss: 0.0017247985815629363
step: 180, loss: 0.03563663735985756
step: 190, loss: 0.0030629276297986507
step: 200, loss: 0.020843736827373505
step: 210, loss: 0.17349594831466675
/usr/local/lib/python3.7/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=0.5363636363636364, f1=0.47880299251870323,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01775752753019333
step: 10, loss: 0.0028229926247149706
step: 20, loss: 0.023453732952475548
step: 30, loss: 0.11238249391317368

```

```

step: 40, loss: 0.028672918677330017
step: 50, loss: 0.001796347671188414
step: 60, loss: 0.009314474649727345
step: 70, loss: 0.05679798871278763
step: 80, loss: 0.006646431051194668
step: 90, loss: 0.06089628487825394
step: 100, loss: 0.002070489339530468
step: 110, loss: 0.0009100447641685605
step: 120, loss: 0.003918597940355539
step: 130, loss: 0.03320206329226494
step: 140, loss: 0.017924247309565544
step: 150, loss: 0.00567997619509697
step: 160, loss: 0.0014368281699717045
step: 170, loss: 0.05034170299768448
step: 180, loss: 0.0013786525232717395
step: 190, loss: 0.012321848422288895
step: 200, loss: 0.004168787505477667
step: 210, loss: 0.04437432438135147
/usr/local/lib/python3.7/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=0.5474613686534215, f1=0.5223529411764705,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028297720476984978
step: 10, loss: 0.018533024936914444
step: 20, loss: 0.013923308812081814
step: 30, loss: 0.008554499596357346
step: 40, loss: 0.007252266630530357
step: 50, loss: 0.002395898336544633
step: 60, loss: 0.0005472752964124084
step: 70, loss: 0.0003738143132068217
step: 80, loss: 0.00013978977221995592
step: 90, loss: 0.0019837385043501854
step: 100, loss: 0.000716279202606529
step: 110, loss: 0.009277504868805408
step: 120, loss: 0.00930145662277937
step: 130, loss: 0.00038174097426235676
step: 140, loss: 0.0006708709406666458
step: 150, loss: 0.006412747781723738
step: 160, loss: 0.000511583813931793
step: 170, loss: 0.007235045079141855
step: 180, loss: 0.0013099689967930317
step: 190, loss: 0.017032772302627563
step: 200, loss: 0.0006061479798518121
step: 210, loss: 0.016441766172647476

```

```

/usr/local/lib/python3.7/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=0.5548387096774194, f1=0.5282167042889391,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006306054070591927
step: 10, loss: 0.007430361118167639
step: 20, loss: 0.026594748720526695
step: 30, loss: 0.004275504034012556
step: 40, loss: 0.0041764178313314915
step: 50, loss: 0.0014047598233446479
step: 60, loss: 0.0012882990995422006
step: 70, loss: 0.0011577195255085826
step: 80, loss: 0.0033380361273884773
step: 90, loss: 0.0032827884424477816
step: 100, loss: 0.002303424058482051
step: 110, loss: 0.003158930689096451
step: 120, loss: 0.0075509920716285706
step: 130, loss: 0.029916463419795036
step: 140, loss: 0.001222072634845972
step: 150, loss: 0.05510585010051727
step: 160, loss: 0.0003065137716475874
step: 170, loss: 0.13608461618423462
step: 180, loss: 0.018931301310658455
step: 190, loss: 0.0009321189136244357
step: 200, loss: 0.0010586394928395748
step: 210, loss: 0.007462369278073311
/usr/local/lib/python3.7/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=0.5522041763341068, f1=0.5000000000000001,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005345013923943043
step: 10, loss: 0.00019209140737075359
step: 20, loss: 0.0020581712014973164
step: 30, loss: 0.00048642989713698626
step: 40, loss: 0.0008666436187922955
step: 50, loss: 0.0030146425124257803
step: 60, loss: 0.004620178136974573
step: 70, loss: 0.000620083708781749
step: 80, loss: 0.001309705781750381
step: 90, loss: 0.0008076686644926667

```

```

step: 100, loss: 0.019377239048480988
step: 110, loss: 0.00039581602322869003
step: 120, loss: 0.0004827367374673486
step: 130, loss: 0.002399857621639967
step: 140, loss: 0.00037493323907256126
step: 150, loss: 0.000140415970236063
step: 160, loss: 0.007665043231099844
step: 170, loss: 0.001002794480882585
step: 180, loss: 0.0007049766718409956
step: 190, loss: 0.009424596093595028
step: 200, loss: 0.010885896161198616
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 210, loss: 0.0005067917518317699
/usr/local/lib/python3.7/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=0.5450450450450449, f1=0.4987775061124694,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0042444257996976376
step: 10, loss: 0.0006816959357820451
step: 20, loss: 0.0004848292446695268
step: 30, loss: 0.0010342855239287019
step: 40, loss: 0.0003780866682063788
step: 50, loss: 0.0002758204354904592
step: 60, loss: 0.0004084737447556108
step: 70, loss: 0.0024510116782039404
step: 80, loss: 0.015424849465489388
step: 90, loss: 0.0003420261782594025
step: 100, loss: 0.02503185346722603
step: 110, loss: 0.00025885325158014894
step: 120, loss: 0.002468985039740801
step: 130, loss: 0.0005520452978089452
step: 140, loss: 0.004146137740463018
step: 150, loss: 0.0006171496934257448
step: 160, loss: 0.00025140075013041496
step: 170, loss: 0.016371598467230797
step: 180, loss: 0.017543788999319077
step: 190, loss: 0.0004785020719282329
step: 200, loss: 0.0037313122302293777
step: 210, loss: 0.005317598581314087
/usr/local/lib/python3.7/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=0.5416666666666667, f1=0.47846889952153104,
best_f1=0.5679012345679013

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006595791201107204
step: 10, loss: 0.00032072816975414753
step: 20, loss: 0.0003727406729012728
step: 30, loss: 0.011335469782352448
step: 40, loss: 0.0012714485637843609
step: 50, loss: 0.019414864480495453
step: 60, loss: 0.0077142962254583836
step: 70, loss: 0.0008592375670559704
step: 80, loss: 0.00991569459438324
step: 90, loss: 0.00328650395385921
step: 100, loss: 0.0002698165480978787
step: 110, loss: 0.005683239549398422
step: 120, loss: 0.0007620083633810282
step: 130, loss: 0.00021269312128424644
step: 140, loss: 0.00017240038141608238
step: 150, loss: 0.001711241900920868
step: 160, loss: 0.0010181355755776167
step: 170, loss: 0.0008090761257335544
step: 180, loss: 0.0009148501558229327
step: 190, loss: 0.001146494410932064
step: 200, loss: 0.00020119927648920566
step: 210, loss: 0.003321868134662509
/usr/local/lib/python3.7/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=0.5485436893203884, f1=0.47355163727959704,
best_f1=0.5679012345679013
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032634756644256413
step: 10, loss: 0.12365321815013885
step: 20, loss: 0.0009384315344505012
step: 30, loss: 0.0002510249032638967
step: 40, loss: 0.012310272082686424
step: 50, loss: 0.0005047020968049765
step: 60, loss: 0.0045846858993172646
step: 70, loss: 0.01222554873675108
step: 80, loss: 0.0004988950095139444
step: 90, loss: 0.0010933129815384746
step: 100, loss: 0.00256793899461627
step: 110, loss: 0.00010920153727056459
step: 120, loss: 0.038380831480026245
step: 130, loss: 0.0017559604020789266
step: 140, loss: 0.00036187778459861875

```

```

step: 150, loss: 0.0009200347121804953
step: 160, loss: 0.0009620571509003639
step: 170, loss: 0.00625198008492589
step: 180, loss: 0.002516739768907428
step: 190, loss: 0.000146712307468988
step: 200, loss: 0.006618126295506954
step: 210, loss: 0.0027048380579799414
/usr/local/lib/python3.7/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=0.5458612975391499, f1=0.491725768321513,
best_f1=0.5679012345679013
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 [00:03, 621.05it/s]
load_f1 = 0.5817490494296578
real_f1 = 0.5730769230769232
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 [00:10, 412.48it/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 distilbert \
  --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 distilbert \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/
```

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

Defaults for this optimization level are:

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

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : O2
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
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: 0.8507311940193176
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.16767218708992004
step: 20, loss: 0.15540331602096558
step: 30, loss: 0.5201919674873352
step: 40, loss: 0.2689386010169983
step: 50, loss: 0.3082183301448822
step: 60, loss: 0.3755294382572174
step: 70, loss: 0.18335314095020294
step: 80, loss: 0.5464164614677429
step: 90, loss: 0.24733883142471313
step: 100, loss: 0.2228717803955078
step: 110, loss: 0.2352716028690338
step: 120, loss: 0.4189845323562622
step: 130, loss: 0.3417733311653137
step: 140, loss: 0.34047234058380127
step: 150, loss: 0.27715107798576355
step: 160, loss: 0.2129170149564743
step: 170, loss: 0.390924870967865
step: 180, loss: 0.2961268126964569
step: 190, loss: 0.14670409262180328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4797979797979798, f1=0.48395061728395067,
best_f1=0.48395061728395067
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29020923376083374
step: 10, loss: 0.026680469512939453
step: 20, loss: 0.1197432205080986
step: 30, loss: 0.2821086645126343
step: 40, loss: 0.5791956186294556
step: 50, loss: 0.3126510679721832
step: 60, loss: 0.2515569031238556
step: 70, loss: 0.2654854357242584
step: 80, loss: 0.17821502685546875
step: 90, loss: 0.15400901436805725
step: 100, loss: 0.24732567369937897
step: 110, loss: 0.2643384337425232
step: 120, loss: 0.21478499472141266
step: 130, loss: 0.25057193636894226
step: 140, loss: 0.2238990068435669
step: 150, loss: 0.07965145260095596
step: 160, loss: 0.10535796731710434
step: 170, loss: 0.2387312799692154
step: 180, loss: 0.21828186511993408
step: 190, loss: 0.17381292581558228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7207207207207206, f1=0.7289156626506026,
best_f1=0.7289156626506026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1951671689748764
step: 10, loss: 0.14026673138141632
step: 20, loss: 0.029321899637579918
step: 30, loss: 0.10986818373203278
step: 40, loss: 0.07597464323043823

```

```

step: 50, loss: 0.24389402568340302
step: 60, loss: 0.09493061900138855
step: 70, loss: 0.134550079703331
step: 80, loss: 0.09314654767513275
step: 90, loss: 0.05300302058458328
step: 100, loss: 0.08321692794561386
step: 110, loss: 0.2049511969089508
step: 120, loss: 0.04707922786474228
step: 130, loss: 0.06005708873271942
step: 140, loss: 0.0417470745742321
step: 150, loss: 0.06966406106948853
step: 160, loss: 0.1080639511346817
step: 170, loss: 0.0325440876185894
step: 180, loss: 0.05277969688177109
step: 190, loss: 0.23186495900154114
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7287671232876712, f1=0.7176781002638523,
best_f1=0.7176781002638523
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06306073069572449
step: 10, loss: 0.06247439235448837
step: 20, loss: 0.03271123021841049
step: 30, loss: 0.040835749357938766
step: 40, loss: 0.011842711828649044
step: 50, loss: 0.007803936023265123
step: 60, loss: 0.04695482552051544
step: 70, loss: 0.0831429734826088
step: 80, loss: 0.03317069634795189
step: 90, loss: 0.05460461974143982
step: 100, loss: 0.04976412653923035
step: 110, loss: 0.07904867827892303
step: 120, loss: 0.0362333320081234
step: 130, loss: 0.17536045610904694
step: 140, loss: 0.022791186347603798
step: 150, loss: 0.05750075355172157
step: 160, loss: 0.14220400154590607
step: 170, loss: 0.008529102429747581
step: 180, loss: 0.15934830904006958
step: 190, loss: 0.037294816225767136
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7380281690140845, f1=0.7569060773480663,
best_f1=0.7569060773480663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009363186545670033
step: 10, loss: 0.009937766939401627
step: 20, loss: 0.07890394330024719
step: 30, loss: 0.015768542885780334
step: 40, loss: 0.04885701835155487
step: 50, loss: 0.034768298268318176
step: 60, loss: 0.005602413322776556
step: 70, loss: 0.0020220214501023293
step: 80, loss: 0.05460815876722336
step: 90, loss: 0.03499109297990799
step: 100, loss: 0.1395179182291031
step: 110, loss: 0.018206778913736343
step: 120, loss: 0.003162194276228547
step: 130, loss: 0.006455365102738142
step: 140, loss: 0.005414399318397045
step: 150, loss: 0.09661129117012024
step: 160, loss: 0.05844588205218315
step: 170, loss: 0.021637016907334328
step: 180, loss: 0.144208624958992
step: 190, loss: 0.16936811804771423
/usr/local/lib/python3.7/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=0.7310704960835509, f1=0.7525773195876289,
best_f1=0.7569060773480663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03596863895654678
step: 10, loss: 0.0036987843923270702
step: 20, loss: 0.0036844101268798113
step: 30, loss: 0.0025745888706296682
step: 40, loss: 0.0035980474203824997
step: 50, loss: 0.020692409947514534
step: 60, loss: 0.0028468414675444365
step: 70, loss: 0.05460519343614578
step: 80, loss: 0.015568161383271217

```

```

step: 90, loss: 0.0032071457244455814
step: 100, loss: 0.020475653931498528
step: 110, loss: 0.009313954040408134
step: 120, loss: 0.04856192693114281
step: 130, loss: 0.0013277132529765368
step: 140, loss: 0.015688581392169
step: 150, loss: 0.07932261377573013
step: 160, loss: 0.0034200563095510006
step: 170, loss: 0.010601788759231567
step: 180, loss: 0.011796976439654827
step: 190, loss: 0.010299825109541416
/usr/local/lib/python3.7/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=0.7071240105540898, f1=0.7282051282051282,
best_f1=0.7569060773480663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016286757308989763
step: 10, loss: 0.006694497540593147
step: 20, loss: 0.002391295274719596
step: 30, loss: 0.025577081367373466
step: 40, loss: 0.05821343883872032
step: 50, loss: 0.016698680818080902
step: 60, loss: 0.01553533785045147
step: 70, loss: 0.0022603184916079044
step: 80, loss: 0.01504630409181118
step: 90, loss: 0.0006875498220324516
step: 100, loss: 0.010868832468986511
step: 110, loss: 0.002602329943329096
step: 120, loss: 0.013261017389595509
step: 130, loss: 0.0015908837085589767
step: 140, loss: 0.018487904220819473
step: 150, loss: 0.0012910813093185425
step: 160, loss: 0.0865061953663826
step: 170, loss: 0.0008262045448645949
step: 180, loss: 0.004052884876728058
step: 190, loss: 0.1167859435081482
/usr/local/lib/python3.7/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=0.7197802197802198, f1=0.7667560321715818,
best_f1=0.7569060773480663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008507254533469677

```

```
step: 10, loss: 0.0008564757299609482
step: 20, loss: 0.002886294387280941
step: 30, loss: 0.002385016530752182
step: 40, loss: 0.011291720904409885
step: 50, loss: 0.0025061534252017736
step: 60, loss: 0.00044116860954090953
step: 70, loss: 0.0023167498875409365
step: 80, loss: 0.009471526369452477
step: 90, loss: 0.0009949051309376955
step: 100, loss: 0.06699928641319275
step: 110, loss: 0.0014469383750110865
step: 120, loss: 0.00508233904838562
step: 130, loss: 0.002084177453070879
step: 140, loss: 0.01496820617467165
step: 150, loss: 0.0012082038447260857
step: 160, loss: 0.00035141114494763315
step: 170, loss: 0.0042263236828148365
step: 180, loss: 0.007810068782418966
step: 190, loss: 0.0028735101222991943
/usr/local/lib/python3.7/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=0.7184466019417476, f1=0.7184466019417476,
best_f1=0.7569060773480663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005980785936117172
step: 10, loss: 0.0016749576898291707
step: 20, loss: 0.024999182671308517
step: 30, loss: 0.0020000338554382324
step: 40, loss: 0.005181108135730028
step: 50, loss: 0.0020402681548148394
step: 60, loss: 0.0021040865685790777
step: 70, loss: 0.0011758487671613693
step: 80, loss: 0.0006835729000158608
step: 90, loss: 0.005804657936096191
step: 100, loss: 0.0006642184453085065
step: 110, loss: 0.0008649916271679103
step: 120, loss: 0.16094771027565002
step: 130, loss: 0.002604277106001973
step: 140, loss: 0.0012443248415365815
step: 150, loss: 0.001985101029276848
step: 160, loss: 0.0013450199039652944
step: 170, loss: 0.009606976062059402
step: 180, loss: 0.012243430130183697
step: 190, loss: 0.0005804484826512635
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7449856733524355, f1=0.7605633802816902,
best_f1=0.7605633802816902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005584273603744805
step: 10, loss: 0.012600700370967388
step: 20, loss: 0.0012249762658029795
step: 30, loss: 0.0011079598916694522
step: 40, loss: 0.00047187888412736356
step: 50, loss: 0.001607623533345759
step: 60, loss: 0.004018635023385286
step: 70, loss: 0.0030159163288772106
step: 80, loss: 0.0002129275817424059
step: 90, loss: 0.01448588352650404
step: 100, loss: 0.0014805716928094625
step: 110, loss: 0.006531424354761839
step: 120, loss: 0.017673851922154427
step: 130, loss: 0.0009887516498565674
step: 140, loss: 0.00033138226717710495
step: 150, loss: 0.0041815913282334805
step: 160, loss: 0.028310516849160194
step: 170, loss: 0.0008325440576300025
step: 180, loss: 0.0004599043168127537
step: 190, loss: 0.0009358433890156448
/usr/local/lib/python3.7/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=0.7431693989071039, f1=0.7513513513513514,
best_f1=0.7605633802816902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018715456826612353
step: 10, loss: 0.0008610811200924218
step: 20, loss: 0.0025603442918509245
step: 30, loss: 0.003687763586640358
step: 40, loss: 0.0002919647376984358
step: 50, loss: 0.0004195026704110205
step: 60, loss: 0.0004929443239234388
step: 70, loss: 0.003083778778091073
step: 80, loss: 0.00028731999918818474

```



```

step: 90, loss: 0.006415359675884247
step: 100, loss: 0.0003683416871353984
step: 110, loss: 0.0003351198974996805
step: 120, loss: 0.00017355731688439846
step: 130, loss: 0.00022087650722824037
step: 140, loss: 0.00040667448774911463
step: 150, loss: 0.0011802598601207137
step: 160, loss: 0.00024101730377878994
step: 170, loss: 0.00030867892201058567
step: 180, loss: 0.0047220936976373196
step: 190, loss: 0.0011057534720748663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7493112947658401, f1=0.7726027397260273,
best_f1=0.7726027397260273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01769060455262661
step: 10, loss: 0.022174054756760597
step: 20, loss: 0.00048385505215264857
step: 30, loss: 0.00186459394171834
step: 40, loss: 0.026712704449892044
step: 50, loss: 0.0019385498017072678
step: 60, loss: 0.0004272785736247897
step: 70, loss: 0.0008990817004814744
step: 80, loss: 0.00031783507438376546
step: 90, loss: 0.0015880962600931525
step: 100, loss: 0.0003101147012785077
step: 110, loss: 0.0003518915909808129
step: 120, loss: 0.0004415885196067393
step: 130, loss: 0.0035953212063759565
step: 140, loss: 0.0004107785935048014
step: 150, loss: 0.00031385626061819494
step: 160, loss: 0.0062427818775177
step: 170, loss: 0.05952077731490135
step: 180, loss: 0.0004229562182445079
step: 190, loss: 0.020972885191440582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7542857142857143, f1=0.7768595041322315,
best_f1=0.7768595041322315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00111129111610353
step: 10, loss: 0.00031873746775090694
step: 20, loss: 0.0005432670004665852
step: 30, loss: 0.0005584544851444662
step: 40, loss: 0.00016702587890904397
step: 50, loss: 0.0012183553772047162
step: 60, loss: 0.0003815061936620623
step: 70, loss: 0.004488300532102585
step: 80, loss: 0.0012068672804161906
step: 90, loss: 0.000379483331926167
step: 100, loss: 0.0005370248109102249
step: 110, loss: 0.0009195051388815045
step: 120, loss: 0.0004448368272278458
step: 130, loss: 0.04021572694182396
step: 140, loss: 0.0010300433496013284
step: 150, loss: 0.0005166113842278719
step: 160, loss: 0.007499131374061108
step: 170, loss: 0.0018099892186000943
step: 180, loss: 0.0040923794731497765
step: 190, loss: 0.00027742423117160797
/usr/local/lib/python3.7/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=0.7535410764872521, f1=0.7616438356164383,
best_f1=0.7768595041322315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00648583984002471
step: 10, loss: 0.002179685514420271
step: 20, loss: 0.0004791701212525368
step: 30, loss: 0.019085364416241646
step: 40, loss: 0.0015631393762305379
step: 50, loss: 0.00042043905705213547
step: 60, loss: 0.0006181774660944939
step: 70, loss: 0.0004258639819454402
step: 80, loss: 0.00017528967873658985
step: 90, loss: 0.00019179977243766189
step: 100, loss: 0.00024002135614864528
step: 110, loss: 0.0005736939492635429
step: 120, loss: 0.00017209938960149884

```

```

step: 130, loss: 0.0002774159947875887
step: 140, loss: 0.0005396465421654284
step: 150, loss: 0.0023375926539301872
step: 160, loss: 0.008193318732082844
step: 170, loss: 0.00025981254293583333
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.00036632470437325537
step: 190, loss: 0.0001759066217346117
/usr/local/lib/python3.7/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=0.7465181058495822, f1=0.766304347826087,
best_f1=0.7768595041322315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002379269280936569
step: 10, loss: 0.0001704921160126105
step: 20, loss: 0.00031121415668167174
step: 30, loss: 0.002710576169192791
step: 40, loss: 0.00036756310146301985
step: 50, loss: 0.0012984283966943622
step: 60, loss: 0.0003519961319398135
step: 70, loss: 0.0004797594156116247
step: 80, loss: 0.00046894288971088827
step: 90, loss: 0.0005431824829429388
step: 100, loss: 0.0002420660457573831
step: 110, loss: 0.00019934594456572086
step: 120, loss: 0.00036311167059466243
step: 130, loss: 0.000498733134008944
step: 140, loss: 0.001419010921381414
step: 150, loss: 0.00042191025568172336
step: 160, loss: 0.0017999312840402126
step: 170, loss: 0.0005360267241485417
step: 180, loss: 0.12418341636657715
step: 190, loss: 0.008272492326796055
/usr/local/lib/python3.7/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=0.7479224376731302, f1=0.766304347826087,
best_f1=0.7768595041322315
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 [00:05, 347.08it/s]
load_f1 = 0.6589861751152073
real_f1 = 0.6436781609195402
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 [00:10, 414.11it/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 distilbert \
--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 distilbert \
  --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: 0.8537509441375732

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

step: 10, loss: 0.22323304414749146

step: 20, loss: 0.15536963939666748

step: 30, loss: 0.24513345956802368

step: 40, loss: 0.31189441680908203

step: 50, loss: 0.37371477484703064

step: 60, loss: 0.43825048208236694

step: 70, loss: 0.3171113133430481

step: 80, loss: 0.24828147888183594

step: 90, loss: 0.40674635767936707

step: 100, loss: 0.22988535463809967

step: 110, loss: 0.18137226998806

step: 120, loss: 0.5600587129592896

step: 130, loss: 0.4079720973968506

step: 140, loss: 0.47182968258857727

step: 150, loss: 0.12017013877630234

step: 160, loss: 0.3392657935619354

step: 170, loss: 0.19392067193984985

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

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

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

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=0.523972602739726, f1=0.5125628140703518,

best_f1=0.5125628140703518

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

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

step: 0, loss: 0.29629650712013245

step: 10, loss: 0.10599435120820999

step: 20, loss: 0.2617230713367462

step: 30, loss: 0.12741297483444214

step: 40, loss: 0.09798905998468399

step: 50, loss: 0.13504955172538757

step: 60, loss: 0.11406174302101135

```

step: 70, loss: 0.17941507697105408
step: 80, loss: 0.1456240862607956
step: 90, loss: 0.21662171185016632
step: 100, loss: 0.09606568515300751
step: 110, loss: 0.28095412254333496
step: 120, loss: 0.1279691904783249
step: 130, loss: 0.10470474511384964
step: 140, loss: 0.2844467759132385
step: 150, loss: 0.12745772302150726
step: 160, loss: 0.1338339000940323
step: 170, loss: 0.05514546111226082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7465437788018434, f1=0.7276785714285715,
best_f1=0.7276785714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08384566009044647
step: 10, loss: 0.10684709250926971
step: 20, loss: 0.015433953143656254
step: 30, loss: 0.20947660505771637
step: 40, loss: 0.008870042860507965
step: 50, loss: 0.10396263003349304
step: 60, loss: 0.21764680743217468
step: 70, loss: 0.20152218639850616
step: 80, loss: 0.20805655419826508
step: 90, loss: 0.11919281631708145
step: 100, loss: 0.0672483816742897
step: 110, loss: 0.15841029584407806
step: 120, loss: 0.005990295670926571
step: 130, loss: 0.16144682466983795
step: 140, loss: 0.009441685862839222
step: 150, loss: 0.054560303688049316
step: 160, loss: 0.09381669759750366
step: 170, loss: 0.13167975842952728
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7499999999999999, f1=0.7663551401869159,
best_f1=0.7663551401869159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08334029465913773
step: 10, loss: 0.07565831393003464
step: 20, loss: 0.018972361460328102
step: 30, loss: 0.05208256468176842
step: 40, loss: 0.01602497324347496
step: 50, loss: 0.07022672146558762
step: 60, loss: 0.054717518389225006
step: 70, loss: 0.0038608054164797068
step: 80, loss: 0.016658222302794456
step: 90, loss: 0.014808663167059422
step: 100, loss: 0.009872804395854473
step: 110, loss: 0.030803564935922623
step: 120, loss: 0.13034658133983612
step: 130, loss: 0.12534894049167633
step: 140, loss: 0.029168499633669853
step: 150, loss: 0.019330831244587898
step: 160, loss: 0.08536075800657272
step: 170, loss: 0.00881549995392561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7540106951871658, f1=0.7355163727959698,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043028779327869415
step: 10, loss: 0.03416166827082634
step: 20, loss: 0.09519259631633759
step: 30, loss: 0.03201600909233093
step: 40, loss: 0.058020904660224915
step: 50, loss: 0.03355148062109947
step: 60, loss: 0.00112847366835922
step: 70, loss: 0.0048549585044384
step: 80, loss: 0.00692548556253314
step: 90, loss: 0.009844970889389515
step: 100, loss: 0.030012313276529312
step: 110, loss: 0.04742470756173134
step: 120, loss: 0.052789296954870224

```



```

step: 130, loss: 0.016729755327105522
step: 140, loss: 0.026120750233530998
step: 150, loss: 0.012277609668672085
step: 160, loss: 0.0812741294503212
step: 170, loss: 0.018365344032645226
/usr/local/lib/python3.7/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=0.7463414634146343, f1=0.7238979118329466,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03873733803629875
step: 10, loss: 0.07456012815237045
step: 20, loss: 0.00395952770486474
step: 30, loss: 0.028605179861187935
step: 40, loss: 0.008264622651040554
step: 50, loss: 0.07142944633960724
step: 60, loss: 0.042430244386196136
step: 70, loss: 0.006069495342671871
step: 80, loss: 0.04682169854640961
step: 90, loss: 0.21180596947669983
step: 100, loss: 0.1299656331539154
step: 110, loss: 0.030442455783486366
step: 120, loss: 0.06892281025648117
step: 130, loss: 0.10703171789646149
step: 140, loss: 0.02122790738940239
step: 150, loss: 0.1264345645904541
step: 160, loss: 0.024553125724196434
step: 170, loss: 0.01628117635846138
/usr/local/lib/python3.7/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=0.7480916030534351, f1=0.7518072289156627,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030742520466446877
step: 10, loss: 0.002292792545631528
step: 20, loss: 0.011645495891571045
step: 30, loss: 0.001160139450803399
step: 40, loss: 0.004794294014573097
step: 50, loss: 0.0037851545494049788
step: 60, loss: 0.22883513569831848
step: 70, loss: 0.04252614080905914
step: 80, loss: 0.0019123853417113423

```

```

step: 90, loss: 0.01306936051696539
step: 100, loss: 0.0033951543737202883
step: 110, loss: 0.002662572544068098
step: 120, loss: 0.11586050689220428
step: 130, loss: 0.05713731423020363
step: 140, loss: 0.0235313568264246
step: 150, loss: 0.14879828691482544
step: 160, loss: 0.006285118870437145
step: 170, loss: 0.10798384249210358
/usr/local/lib/python3.7/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=0.7409200968523002, f1=0.7425968109339408,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018403470516204834
step: 10, loss: 0.001738921389915049
step: 20, loss: 0.03286813199520111
step: 30, loss: 0.01331913098692894
step: 40, loss: 0.0009200367494486272
step: 50, loss: 0.000589504896197468
step: 60, loss: 0.008969902992248535
step: 70, loss: 0.010417837649583817
step: 80, loss: 0.006689118687063456
step: 90, loss: 0.010170831345021725
step: 100, loss: 0.0038770854007452726
step: 110, loss: 0.004143958445638418
step: 120, loss: 0.0014966446906328201
step: 130, loss: 0.0004015110316686332
step: 140, loss: 0.10348614305257797
step: 150, loss: 0.10939370840787888
step: 160, loss: 0.036220353096723557
step: 170, loss: 0.003310595639050007
/usr/local/lib/python3.7/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=0.753623188405797, f1=0.7494145199063231,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005969488061964512
step: 10, loss: 0.007269591558724642
step: 20, loss: 0.0002826797717716545
step: 30, loss: 0.05586151406168938
step: 40, loss: 0.017344659194350243

```

```

step: 50, loss: 0.014118772931396961
step: 60, loss: 0.000513987906742841
step: 70, loss: 0.026461808010935783
step: 80, loss: 0.018138542771339417
step: 90, loss: 0.002245643176138401
step: 100, loss: 0.010367234237492085
step: 110, loss: 0.00021096992713864893
step: 120, loss: 0.014465628191828728
step: 130, loss: 0.1655881106853485
step: 140, loss: 0.0003525992506183684
step: 150, loss: 0.012102550826966763
step: 160, loss: 0.08116939663887024
step: 170, loss: 0.013050812296569347
/usr/local/lib/python3.7/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=0.7462686567164178, f1=0.7417840375586854,
best_f1=0.7355163727959698
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07931183278560638
step: 10, loss: 0.01227493304759264
step: 20, loss: 0.001284275553189218
step: 30, loss: 0.0005388561985455453
step: 40, loss: 0.004045614041388035
step: 50, loss: 0.0005849784938618541
step: 60, loss: 0.004268226213753223
step: 70, loss: 0.0031722518615424633
step: 80, loss: 0.012872313149273396
step: 90, loss: 0.04332214966416359
step: 100, loss: 0.002707804087549448
step: 110, loss: 0.0006177463219501078
step: 120, loss: 0.02210608497262001
step: 130, loss: 0.01100879069417715
step: 140, loss: 0.014790275134146214
step: 150, loss: 0.002688647946342826
step: 160, loss: 0.0007595170172862709
step: 170, loss: 0.00034889738890342414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7567567567567567, f1=0.7499999999999999,
best_f1=0.7499999999999999

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027638705796562135
step: 10, loss: 0.0066664800979197025
step: 20, loss: 0.03237401694059372
step: 30, loss: 0.009195924736559391
step: 40, loss: 0.006267141550779343
step: 50, loss: 0.0009325784631073475
step: 60, loss: 0.003555922070518136
step: 70, loss: 0.020860858261585236
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.00018245624960400164
step: 90, loss: 0.0002666029322426766
step: 100, loss: 0.0001784577325452119
step: 110, loss: 0.0007170604076236486
step: 120, loss: 0.0010256008245050907
step: 130, loss: 0.0032976665534079075
step: 140, loss: 0.021110856905579567
step: 150, loss: 0.002288123592734337
step: 160, loss: 0.013312317430973053
step: 170, loss: 0.14449678361415863
/usr/local/lib/python3.7/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=0.7360406091370559, f1=0.7338129496402879,
best_f1=0.7499999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00489921635016799
step: 10, loss: 0.06559756398200989
step: 20, loss: 0.03808052837848663
step: 30, loss: 0.0006140075856819749
step: 40, loss: 0.01398796122521162
step: 50, loss: 0.0001783047046046704
step: 60, loss: 0.0003486594941932708
step: 70, loss: 0.013180317357182503
step: 80, loss: 0.03897128254175186
step: 90, loss: 0.001066838507540524
step: 100, loss: 0.0007297797128558159
step: 110, loss: 0.00015956288552843034
step: 120, loss: 0.0007830106769688427
step: 130, loss: 0.06431598961353302
step: 140, loss: 0.0018200704362243414
step: 150, loss: 0.0011221899185329676
step: 160, loss: 0.0017689191736280918
step: 170, loss: 0.0006571048870682716

```

```

/usr/local/lib/python3.7/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=0.7407407407407407, f1=0.7308641975308643,
best_f1=0.7499999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017126263992395252
step: 10, loss: 0.04222218319773674
step: 20, loss: 0.0007220800616778433
step: 30, loss: 0.0009025735198520124
step: 40, loss: 0.00014587075565941632
step: 50, loss: 0.00018763371917884797
step: 60, loss: 0.00030018642428331077
step: 70, loss: 0.01959548331797123
step: 80, loss: 0.0050080763176083565
step: 90, loss: 0.0004292766097933054
step: 100, loss: 0.0001583188131917268
step: 110, loss: 0.001087732263840735
step: 120, loss: 0.0012732745381072164
step: 130, loss: 0.0008527339086867869
step: 140, loss: 0.00012874948151875287
step: 150, loss: 0.001908609177917242
step: 160, loss: 0.001846618251875043
step: 170, loss: 0.007049859501421452
/usr/local/lib/python3.7/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=0.7526315789473685, f1=0.7487684729064039,
best_f1=0.7499999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034387510269880295
step: 10, loss: 0.03151153400540352
step: 20, loss: 0.011787893250584602
step: 30, loss: 0.0014565751189365983
step: 40, loss: 0.0005072802305221558
step: 50, loss: 0.003127056173980236
step: 60, loss: 0.00022401282330974936
step: 70, loss: 0.00280662695877254
step: 80, loss: 0.0007748089265078306
step: 90, loss: 0.000782441406045109
step: 100, loss: 0.00032669640495441854
step: 110, loss: 0.00041328175575472414
step: 120, loss: 0.002821736503392458
step: 130, loss: 0.0004877594765275717

```

```

step: 140, loss: 0.0012797341914847493
step: 150, loss: 0.003897444810718298
step: 160, loss: 0.04971832409501076
step: 170, loss: 0.0002644648193381727
/usr/local/lib/python3.7/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=0.7540106951871658, f1=0.7444168734491315,
best_f1=0.7499999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005741436034440994
step: 10, loss: 0.004679146222770214
step: 20, loss: 0.019982971251010895
step: 30, loss: 0.012675515376031399
step: 40, loss: 0.0001928116980707273
step: 50, loss: 0.00795521680265665
step: 60, loss: 0.00015379003889393061
step: 70, loss: 0.03471347689628601
step: 80, loss: 0.0005616985727101564
step: 90, loss: 0.00017099647084251046
step: 100, loss: 0.004203770309686661
step: 110, loss: 0.0002688643289729953
step: 120, loss: 0.023446833714842796
step: 130, loss: 0.0017370181158185005
step: 140, loss: 0.004723379388451576
step: 150, loss: 0.0024365317076444626
step: 160, loss: 0.0001470986899221316
step: 170, loss: 0.00025222255499102175
/usr/local/lib/python3.7/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=0.751918158567775, f1=0.7565011820330969,
best_f1=0.7499999999999999
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 [00:04, 448.78it/s]

load_f1 = 0.6893424036281178

real_f1 = 0.6772009029345374

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 [00:10, 412.63it/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 distilbert \
  --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 distilbert \
  --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: 0.7993522882461548
```

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

```
step: 10, loss: 0.490243136882782
```

```
step: 20, loss: 0.6174472570419312
```

```
step: 30, loss: 0.49588701128959656
```

```
step: 40, loss: 0.359375923871994
```

```
step: 50, loss: 0.2573367655277252
```

```
step: 60, loss: 0.19991044700145721
```

```
step: 70, loss: 0.2893015742301941
```

```
step: 80, loss: 0.24645887315273285
```

```
step: 90, loss: 0.14010383188724518
```

```
step: 100, loss: 0.2089306265115738
```

```
step: 110, loss: 0.12750115990638733
```

```
step: 120, loss: 0.04263770952820778
```

```
step: 130, loss: 0.01952025666832924
```

```
step: 140, loss: 0.06442736089229584
```

```
step: 150, loss: 0.15456415712833405
```

```
step: 160, loss: 0.3177449405193329
```

```
step: 170, loss: 0.02694408781826496
```

```
step: 180, loss: 0.004682376980781555
```

```
step: 190, loss: 0.07881078869104385
```

```
step: 200, loss: 0.02766128070652485
```

```
step: 210, loss: 0.03197872266173363
```

```
step: 220, loss: 0.010255422443151474
```

```
step: 230, loss: 0.17711463570594788
```

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

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

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

```
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=0.9568106312292359, f1=0.9485458612975392,
```

```
best_f1=0.9485458612975392
```

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

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

```
step: 0, loss: 0.07659543305635452
```

```
step: 10, loss: 0.009947694838047028
```

```
step: 20, loss: 0.03262772038578987
```

```
step: 30, loss: 0.009702326729893684
```

```
step: 40, loss: 0.010536318644881248
```

```
step: 50, loss: 0.07312121987342834
```

```
step: 60, loss: 0.14363490045070648
```

```

step: 70, loss: 0.010653618723154068
step: 80, loss: 0.00824529305100441
step: 90, loss: 0.010958747938275337
step: 100, loss: 0.15842010080814362
step: 110, loss: 0.10548511147499084
step: 120, loss: 0.016477398574352264
step: 130, loss: 0.019230110570788383
step: 140, loss: 0.07480127364397049
step: 150, loss: 0.08719097822904587
step: 160, loss: 0.015161282382905483
step: 170, loss: 0.0725138783454895
step: 180, loss: 0.0248437337577343
step: 190, loss: 0.1217796802520752
step: 200, loss: 0.07716267555952072
step: 210, loss: 0.10065476596355438
step: 220, loss: 0.0017684687627479434
step: 230, loss: 0.018688252195715904
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9590254706533776, f1=0.9535398230088497,
best_f1=0.9535398230088497
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0583559088408947
step: 10, loss: 0.007499130442738533
step: 20, loss: 0.01415508333593607
step: 30, loss: 0.006936630699783564
step: 40, loss: 0.22692792117595673
step: 50, loss: 0.006026057060807943
step: 60, loss: 0.05007009208202362
step: 70, loss: 0.008264435455203056
step: 80, loss: 0.005957588087767363
step: 90, loss: 0.043176889419555664
step: 100, loss: 0.009641948156058788
step: 110, loss: 0.06346620619297028
step: 120, loss: 0.015359602868556976
step: 130, loss: 0.017964942380785942
step: 140, loss: 0.004347005393356085
step: 150, loss: 0.0019510382553562522
step: 160, loss: 0.03699323907494545
step: 170, loss: 0.01493852399289608
step: 180, loss: 0.0024591307155787945

```

```

step: 190, loss: 0.02769802138209343
step: 200, loss: 0.014962850138545036
step: 210, loss: 0.030335651710629463
step: 220, loss: 0.01019643247127533
step: 230, loss: 0.036805786192417145
/usr/local/lib/python3.7/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=0.9511111111111112, f1=0.9403973509933775,
best_f1=0.9535398230088497
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020330015569925308
step: 10, loss: 0.012859180569648743
step: 20, loss: 0.0029051597230136395
step: 30, loss: 0.052365511655807495
step: 40, loss: 0.02007168158888817
step: 50, loss: 0.00204737507738173
step: 60, loss: 0.0043253409676253796
step: 70, loss: 0.10368962585926056
step: 80, loss: 0.15729881823062897
step: 90, loss: 0.10129525512456894
step: 100, loss: 0.01284282561391592
step: 110, loss: 0.0312163345515728
step: 120, loss: 0.006782342679798603
step: 130, loss: 0.003771938383579254
step: 140, loss: 0.005549309309571981
step: 150, loss: 0.007458563894033432
step: 160, loss: 0.0019665074069052935
step: 170, loss: 0.030565129593014717
step: 180, loss: 0.10277878493070602
step: 190, loss: 0.007296170108020306
step: 200, loss: 0.005739601794630289
step: 210, loss: 0.0051632234826684
step: 220, loss: 0.002972689690068364
step: 230, loss: 0.003981580026447773
/usr/local/lib/python3.7/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=0.9478260869565216, f1=0.9485213581599123,
best_f1=0.9535398230088497
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012715510092675686
step: 10, loss: 0.023906728252768517
step: 20, loss: 0.0033568127546459436

```

```

step: 30, loss: 0.0008317931205965579
step: 40, loss: 0.00048578408313915133
step: 50, loss: 0.0004162382974755019
step: 60, loss: 0.0006040802109055221
step: 70, loss: 0.0012877568369731307
step: 80, loss: 0.000836987397633493
step: 90, loss: 0.005487927235662937
step: 100, loss: 0.019348440691828728
step: 110, loss: 0.0019680941477417946
step: 120, loss: 0.049368590116500854
step: 130, loss: 0.006106330547481775
step: 140, loss: 0.0015089758671820164
step: 150, loss: 0.0021396735683083534
step: 160, loss: 0.024096999317407608
step: 170, loss: 0.0530778132379055
step: 180, loss: 0.001022370532155037
step: 190, loss: 0.0007788373623043299
step: 200, loss: 0.0023431573063135147
step: 210, loss: 0.0007797859143465757
step: 220, loss: 0.001019371789880097
step: 230, loss: 0.011252518743276596
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.964835164835165, f1=0.9537444933920706,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028941105119884014
step: 10, loss: 0.02758517861366272
step: 20, loss: 0.0045486316084861755
step: 30, loss: 0.00041872504516504705
step: 40, loss: 0.0020508170127868652
step: 50, loss: 0.0021340635139495134
step: 60, loss: 0.0007091641309671104
step: 70, loss: 0.0015870489878579974
step: 80, loss: 0.04115710407495499
step: 90, loss: 0.00044096793862991035
step: 100, loss: 0.0025819106958806515
step: 110, loss: 0.0287519209086895
step: 120, loss: 0.0003587734536267817
step: 130, loss: 0.00023597717517986894
step: 140, loss: 0.000648858433123678

```

```

step: 150, loss: 0.17609310150146484
step: 160, loss: 0.06734465807676315
step: 170, loss: 0.0023699041921645403
step: 180, loss: 0.000661147350911051
step: 190, loss: 0.0008095522061921656
step: 200, loss: 0.014091972261667252
step: 210, loss: 0.0030267315451055765
step: 220, loss: 0.0007519870414398611
step: 230, loss: 0.0002100977289956063
/usr/local/lib/python3.7/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=0.9632925472747497, f1=0.9498327759197325,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04437752440571785
step: 10, loss: 0.00039894055225886405
step: 20, loss: 0.002050269627943635
step: 30, loss: 0.0006160876364447176
step: 40, loss: 0.003498827340081334
step: 50, loss: 0.00019265587616246194
step: 60, loss: 0.0009697307832539082
step: 70, loss: 0.006103395950049162
step: 80, loss: 0.0006335412035696208
step: 90, loss: 0.0037402433808892965
step: 100, loss: 0.0012428424088284373
step: 110, loss: 0.0005304781370796263
step: 120, loss: 0.0014318187022581697
step: 130, loss: 0.0007836418808437884
step: 140, loss: 0.00015343763516284525
step: 150, loss: 0.00021059028222225606
step: 160, loss: 0.0052693188190460205
step: 170, loss: 0.00027907558251172304
step: 180, loss: 0.0001660075649851933
step: 190, loss: 0.0010056180180981755
step: 200, loss: 0.0016933941515162587
step: 210, loss: 0.0012693294556811452
step: 220, loss: 0.0002530618803575635
step: 230, loss: 0.004573395941406488
/usr/local/lib/python3.7/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=0.9636963696369637, f1=0.9513274336283186,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027474213391542435
step: 10, loss: 0.033460065722465515
step: 20, loss: 0.0015080795856192708
step: 30, loss: 0.0001168606395367533
step: 40, loss: 0.00015692535089328885
step: 50, loss: 0.0007971416343934834
step: 60, loss: 0.00011835367331514135
step: 70, loss: 0.0002462189586367458
step: 80, loss: 0.005666074343025684
step: 90, loss: 0.00013257811951916665
step: 100, loss: 0.0011517053935676813
step: 110, loss: 0.0001570290478412062
step: 120, loss: 0.00011944361904170364
step: 130, loss: 0.01042563933879137
step: 140, loss: 0.0008229715749621391
step: 150, loss: 8.036979852477089e-05
step: 160, loss: 0.0003551617846824229
step: 170, loss: 8.590721699874848e-05
step: 180, loss: 0.001149150193668902
step: 190, loss: 0.00021573946287389845
step: 200, loss: 0.002868682611733675
step: 210, loss: 8.348791016032919e-05
step: 220, loss: 0.00011903489212272689
step: 230, loss: 0.03554482012987137
/usr/local/lib/python3.7/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=0.9594742606790799, f1=0.9544950055493895,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005628036451525986
step: 10, loss: 0.0010543983662500978
step: 20, loss: 0.0027041416615247726
step: 30, loss: 0.00020241538004484028
step: 40, loss: 7.979032670846209e-05
step: 50, loss: 6.738872616551816e-05
step: 60, loss: 0.0001311331579927355
step: 70, loss: 0.00593353109434247
step: 80, loss: 0.044725071638822556
step: 90, loss: 0.0004002757777925581
step: 100, loss: 0.000160858835442923
step: 110, loss: 8.521154086338356e-05
step: 120, loss: 0.002580981934443116
step: 130, loss: 0.00014766283857170492
step: 140, loss: 0.03791332244873047

```

```

step: 150, loss: 8.920839172787964e-05
step: 160, loss: 0.00035048360587097704
step: 170, loss: 0.00018012491636909544
step: 180, loss: 0.00031000457238405943
step: 190, loss: 0.00035742580075748265
step: 200, loss: 0.00011907418229384348
step: 210, loss: 0.00035179860424250364
step: 220, loss: 0.032919686287641525
step: 230, loss: 0.0006852697115391493
/usr/local/lib/python3.7/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=0.9576837416481068, f1=0.9486607142857143,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001164135173894465
step: 10, loss: 6.212405423866585e-05
step: 20, loss: 0.00047771583194844425
step: 30, loss: 0.0001490594877395779
step: 40, loss: 3.8830945413792506e-05
step: 50, loss: 0.0001541497913422063
step: 60, loss: 0.000572583288885653
step: 70, loss: 0.0001003226061584428
step: 80, loss: 0.0001581735850777477
step: 90, loss: 0.00013594822667073458
step: 100, loss: 0.00014422413369175047
step: 110, loss: 0.007619024254381657
step: 120, loss: 0.022291749715805054
step: 130, loss: 0.006050756201148033
step: 140, loss: 0.003958778455853462
step: 150, loss: 7.665049633942544e-05
step: 160, loss: 5.297092502587475e-05
step: 170, loss: 0.0002059822145383805
step: 180, loss: 0.00012060673179803416
step: 190, loss: 0.00012037286069244146
step: 200, loss: 4.71859602839686e-05
step: 210, loss: 5.730066550313495e-05
step: 220, loss: 0.00011877954239025712
step: 230, loss: 0.00032155579538084567
/usr/local/lib/python3.7/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=0.9601769911504424, f1=0.9513274336283186,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014611140068154782
step: 10, loss: 0.00013535135076381266
step: 20, loss: 0.0007722307927906513
step: 30, loss: 0.006309512071311474
step: 40, loss: 0.04411046952009201
step: 50, loss: 0.0007929742569103837
step: 60, loss: 0.00019444571807980537
step: 70, loss: 0.00013759135617874563
step: 80, loss: 0.00010480541095603257
step: 90, loss: 0.0001077250053640455
step: 100, loss: 0.0002941882994491607
step: 110, loss: 0.0005626280908472836
step: 120, loss: 0.00011925184662686661
step: 130, loss: 6.667948036920279e-05
step: 140, loss: 5.431022145785391e-05
step: 150, loss: 7.978414942044765e-05
step: 160, loss: 5.061558840679936e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.10827518999576569
step: 180, loss: 0.0174521766602993
step: 190, loss: 0.0009180265478789806
step: 200, loss: 0.00019530262215994298
step: 210, loss: 5.8042605814989656e-05
step: 220, loss: 8.852115570334718e-05
step: 230, loss: 0.0021844664588570595
/usr/local/lib/python3.7/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=0.9507119386637459, f1=0.9432314410480349,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021545342169702053
step: 10, loss: 0.00021268974523991346
step: 20, loss: 6.141510675661266e-05
step: 30, loss: 0.00045493923244066536
step: 40, loss: 5.281386256683618e-05
step: 50, loss: 6.383851723512635e-05
step: 60, loss: 0.006196224596351385
step: 70, loss: 0.00012750735913868994
step: 80, loss: 0.00010362762986915186
step: 90, loss: 0.00014077656669542193
step: 100, loss: 0.028935732319951057
step: 110, loss: 0.0007265640888363123
step: 120, loss: 0.0001803016202757135
step: 130, loss: 0.00012792077905032784

```



```

step: 140, loss: 0.00011573448136914521
step: 150, loss: 0.0002539357228670269
step: 160, loss: 0.006372413598001003
step: 170, loss: 6.588427640963346e-05
step: 180, loss: 0.0002619186125230044
step: 190, loss: 7.51204788684845e-05
step: 200, loss: 0.0012195915915071964
step: 210, loss: 8.72768578119576e-05
step: 220, loss: 0.10320664197206497
step: 230, loss: 9.818169928621501e-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=0.9553571428571428, f1=0.9463087248322148,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015506842464674264
step: 10, loss: 0.00022511942370329052
step: 20, loss: 0.00013630058674607426
step: 30, loss: 0.028041645884513855
step: 40, loss: 0.00026306198560632765
step: 50, loss: 0.0002477726084180176
step: 60, loss: 0.00011184602772118524
step: 70, loss: 6.523217598441988e-05
step: 80, loss: 7.964426913531497e-05
step: 90, loss: 5.5977019655983895e-05
step: 100, loss: 0.0001125594208133407
step: 110, loss: 0.0002849239099305123
step: 120, loss: 0.0014455527998507023
step: 130, loss: 0.0001481907966081053
step: 140, loss: 0.00023736983712296933
step: 150, loss: 0.0006221827352419496
step: 160, loss: 0.0012419297127053142
step: 170, loss: 6.908728391863406e-05
step: 180, loss: 0.0001884346129372716
step: 190, loss: 8.901341789169237e-05
step: 200, loss: 0.0005512132192961872
step: 210, loss: 9.602538193576038e-05
step: 220, loss: 0.00069003168027848
step: 230, loss: 7.142496178857982e-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=0.9550930996714129, f1=0.9472527472527472,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.037831917637959e-05
step: 10, loss: 5.912597771384753e-05
step: 20, loss: 0.010459394194185734
step: 30, loss: 8.753919246373698e-05
step: 40, loss: 3.9225506043294445e-05
step: 50, loss: 0.0007013012073002756
step: 60, loss: 0.0002409839944448322
step: 70, loss: 0.0004391803522594273
step: 80, loss: 9.508983930572867e-05
step: 90, loss: 0.0005593578098341823
step: 100, loss: 0.0010364389745518565
step: 110, loss: 0.0022366056218743324
step: 120, loss: 9.426850010640919e-05
step: 130, loss: 6.473838584497571e-05
step: 140, loss: 0.00012352655176073313
step: 150, loss: 0.0004264904127921909
step: 160, loss: 3.1269126338884234e-05
step: 170, loss: 5.612914901576005e-05
step: 180, loss: 5.7347420806763694e-05
step: 190, loss: 5.942392090219073e-05
step: 200, loss: 5.799182690680027e-05
step: 210, loss: 0.0001631239865673706
step: 220, loss: 5.0429291150067e-05
step: 230, loss: 3.2479885703651235e-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=0.958057395143488, f1=0.9513274336283186,
best_f1=0.9537444933920706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.868458563578315e-05
step: 10, loss: 0.0038680557627230883
step: 20, loss: 0.00011718024325091392
step: 30, loss: 0.0027935716789215803
step: 40, loss: 6.218232010724023e-05
step: 50, loss: 7.661989366170019e-05
step: 60, loss: 6.108925299486145e-05
step: 70, loss: 9.58587770583108e-05
step: 80, loss: 0.00022464575886260718
step: 90, loss: 3.947442746721208e-05
step: 100, loss: 4.4884229282615706e-05
step: 110, loss: 4.8005316784838215e-05
step: 120, loss: 5.187174610910006e-05
step: 130, loss: 8.986380998976529e-05

```

```


step: 140, loss: 0.00011276035365881398
step: 150, loss: 7.577292853966355e-05
step: 160, loss: 0.00012266448175068945
step: 170, loss: 5.4765347158536315e-05
step: 180, loss: 4.025735688628629e-05
step: 190, loss: 0.00044761941535398364
step: 200, loss: 5.423451875685714e-05
step: 210, loss: 0.00015422221622429788
step: 220, loss: 0.0002177295828005299
step: 230, loss: 6.34493408142589e-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=0.9570011025358324, f1=0.9492273730684326,
best_f1=0.9537444933920706
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 [00:07, 337.04it/s]
load_f1 = 0.958659217877095
real_f1 = 0.9574944071588367

```

```
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 [00:10, 413.92it/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 distilbert \  
  --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 distilbert \  
  --use_gpu \  
  --fp16 \  
  --checkpoint_path checkpoints/
```

```
Downloading: 100% 442/442 [00:00<00:00, 370kB/s]  
Downloading: 100% 232k/232k [00:00<00:00, 348kB/s]  
Downloading: 100% 268M/268M [00:03<00:00, 71.4MB/s]  
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32  
master weights.
```

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.7914949655532837
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4486654996871948
step: 20, loss: 0.5108965039253235
step: 30, loss: 0.4660613536834717
step: 40, loss: 0.4715266823768616
step: 50, loss: 0.41198021173477173
step: 60, loss: 0.44517576694488525
step: 70, loss: 0.07474343478679657
step: 80, loss: 0.20514564216136932
step: 90, loss: 0.34602048993110657
step: 100, loss: 0.2281731814146042
step: 110, loss: 0.0537647046148777
step: 120, loss: 0.13854485750198364
step: 130, loss: 0.06487289816141129
step: 140, loss: 0.22630351781845093

```

```

step: 150, loss: 0.0792669951915741
step: 160, loss: 0.16117988526821136
step: 170, loss: 0.2853356897830963
step: 180, loss: 0.17421910166740417
step: 190, loss: 0.12226370722055435
step: 200, loss: 0.15815629065036774
step: 210, loss: 0.09665179252624512
step: 220, loss: 0.19013060629367828
step: 230, loss: 0.10877477377653122
step: 240, loss: 0.14248086512088776
step: 250, loss: 0.09535862505435944
step: 260, loss: 0.031544219702482224
step: 270, loss: 0.034270137548446655
step: 280, loss: 0.2913488745689392
step: 290, loss: 0.07639994472265244
step: 300, loss: 0.3286646008491516
step: 310, loss: 0.11639988422393799
step: 320, loss: 0.12484593689441681
step: 330, loss: 0.17033471167087555
step: 340, loss: 0.2366679161787033
step: 350, loss: 0.07291418313980103
step: 360, loss: 0.05463821068406105
step: 370, loss: 0.10286528617143631
step: 380, loss: 0.24167585372924805
step: 390, loss: 0.02785877138376236
step: 400, loss: 0.07445946335792542
step: 410, loss: 0.09928398579359055
step: 420, loss: 0.04012315347790718
step: 430, loss: 0.05182782933115959
step: 440, loss: 0.13555191457271576
step: 450, loss: 0.052769772708415985
step: 460, loss: 0.017478305846452713
step: 470, loss: 0.19549037516117096
step: 480, loss: 0.23811545968055725
step: 490, loss: 0.1338135004043579
step: 500, loss: 0.12237893790006638
step: 510, loss: 0.11932113766670227
step: 520, loss: 0.15483875572681427
step: 530, loss: 0.0580025240778923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9111424541607899, f1=0.9068150208623087,
best_f1=0.9068150208623087

```

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

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

```
step: 0, loss: 0.09295061975717545  
step: 10, loss: 0.13927923142910004  
step: 20, loss: 0.14180594682693481  
step: 30, loss: 0.06649167090654373  
step: 40, loss: 0.011498285457491875  
step: 50, loss: 0.1053260788321495  
step: 60, loss: 0.18705637753009796  
step: 70, loss: 0.28512290120124817  
step: 80, loss: 0.01633468270301819  
step: 90, loss: 0.07245275378227234  
step: 100, loss: 0.2638527750968933  
step: 110, loss: 0.0654277354478836  
step: 120, loss: 0.06494010984897614  
step: 130, loss: 0.04136290401220322  
step: 140, loss: 0.0408177450299263  
step: 150, loss: 0.06355009227991104  
step: 160, loss: 0.014434409327805042  
step: 170, loss: 0.12808340787887573  
step: 180, loss: 0.025984203442931175  
step: 190, loss: 0.03872930258512497  
step: 200, loss: 0.024009035900235176  
step: 210, loss: 0.019848652184009552  
step: 220, loss: 0.15372103452682495  
step: 230, loss: 0.07958108186721802  
step: 240, loss: 0.1475164145231247  
step: 250, loss: 0.1306813806295395  
step: 260, loss: 0.022771533578634262  
step: 270, loss: 0.057509999722242355  
step: 280, loss: 0.09672864526510239  
step: 290, loss: 0.12543578445911407  
step: 300, loss: 0.04747777059674263  
step: 310, loss: 0.06743323802947998  
step: 320, loss: 0.09951122850179672  
step: 330, loss: 0.02230602130293846  
step: 340, loss: 0.004085815045982599  
step: 350, loss: 0.03290483355522156  
step: 360, loss: 0.1015494093298912  
step: 370, loss: 0.02969176322221756  
step: 380, loss: 0.07111489027738571  
step: 390, loss: 0.04208279028534889  
step: 400, loss: 0.22863198816776276  
step: 410, loss: 0.01901191845536232  
step: 420, loss: 0.04423549398779869  
step: 430, loss: 0.04235658794641495  
step: 440, loss: 0.010645493865013123
```

```

step: 450, loss: 0.024282172322273254
step: 460, loss: 0.318903386592865
step: 470, loss: 0.054934047162532806
step: 480, loss: 0.23542770743370056
step: 490, loss: 0.03532914072275162
step: 500, loss: 0.03657149523496628
step: 510, loss: 0.055757179856300354
step: 520, loss: 0.007284918334335089
step: 530, loss: 0.1173316165804863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9179869524697111, f1=0.9094287041337669,
best_f1=0.9094287041337669
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07518612593412399
step: 10, loss: 0.08008851110935211
step: 20, loss: 0.16996508836746216
step: 30, loss: 0.1019519492983818
step: 40, loss: 0.003623948898166418
step: 50, loss: 0.04698282852768898
step: 60, loss: 0.016873857006430626
step: 70, loss: 0.012756521813571453
step: 80, loss: 0.0259856004267931
step: 90, loss: 0.11119674146175385
step: 100, loss: 0.02124462090432644
step: 110, loss: 0.019957253709435463
step: 120, loss: 0.02726522646844387
step: 130, loss: 0.04183325916528702
step: 140, loss: 0.024404244497418404
step: 150, loss: 0.01183006539940834
step: 160, loss: 0.008066464215517044
step: 170, loss: 0.03959217295050621
step: 180, loss: 0.003521672682836652
step: 190, loss: 0.006982248742133379
step: 200, loss: 0.013775226660072803
step: 210, loss: 0.028495769947767258
step: 220, loss: 0.02051364630460739
step: 230, loss: 0.022701285779476166
step: 240, loss: 0.015241798013448715
step: 250, loss: 0.005463649518787861
step: 260, loss: 0.0056473626755177975

```



```

step: 270, loss: 0.030659260228276253
step: 280, loss: 0.03253563866019249
step: 290, loss: 0.03352366015315056
step: 300, loss: 0.14554771780967712
step: 310, loss: 0.07263398170471191
step: 320, loss: 0.265087753534317
step: 330, loss: 0.006355868186801672
step: 340, loss: 0.009217900224030018
step: 350, loss: 0.07983596622943878
step: 360, loss: 0.023052716627717018
step: 370, loss: 0.03680826723575592
step: 380, loss: 0.019158123061060905
step: 390, loss: 0.07449015229940414
step: 400, loss: 0.085858054459095
step: 410, loss: 0.023014072328805923
step: 420, loss: 0.02793688140809536
step: 430, loss: 0.015090499073266983
step: 440, loss: 0.05836353078484535
step: 450, loss: 0.025636833161115646
step: 460, loss: 0.1368456333875656
step: 470, loss: 0.028912674635648727
step: 480, loss: 0.004572204779833555
step: 490, loss: 0.11887817084789276
step: 500, loss: 0.12698616087436676
step: 510, loss: 0.009765085764229298
step: 520, loss: 0.0406239815056324
step: 530, loss: 0.03573695570230484
/usr/local/lib/python3.7/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=0.9132507149666349, f1=0.9090909090909091,
best_f1=0.9094287041337669
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01629490591585636
step: 10, loss: 0.011078206822276115
step: 20, loss: 0.020086202770471573
step: 30, loss: 0.011397662572562695
step: 40, loss: 0.006559322122484446
step: 50, loss: 0.10554737597703934
step: 60, loss: 0.0036801036912947893
step: 70, loss: 0.0021335165947675705
step: 80, loss: 0.025959761813282967
step: 90, loss: 0.007885833270847797
step: 100, loss: 0.002903175773099065
step: 110, loss: 0.014989052899181843
step: 120, loss: 0.0010910200653597713

```

```

step: 130, loss: 0.00608909921720624
step: 140, loss: 0.06359968334436417
step: 150, loss: 0.0009527382790111005
step: 160, loss: 0.007346646394580603
step: 170, loss: 0.039389755576848984
step: 180, loss: 0.006545423995703459
step: 190, loss: 0.05385664477944374
step: 200, loss: 0.040871184319257736
step: 210, loss: 0.005483776796609163
step: 220, loss: 0.010831774212419987
step: 230, loss: 0.22583135962486267
step: 240, loss: 0.007782479282468557
step: 250, loss: 0.0006389959598891437
step: 260, loss: 0.10314332693815231
step: 270, loss: 0.05438768491148949
step: 280, loss: 0.03701247647404671
step: 290, loss: 0.10766424238681793
step: 300, loss: 0.0016863553319126368
step: 310, loss: 0.013685708865523338
step: 320, loss: 0.011625709012150764
step: 330, loss: 0.014284268021583557
step: 340, loss: 0.02059393748641014
step: 350, loss: 0.10648462921380997
step: 360, loss: 0.004061810206621885
step: 370, loss: 0.08137413114309311
step: 380, loss: 0.005910647567361593
step: 390, loss: 0.03739459812641144
step: 400, loss: 0.18672890961170197
step: 410, loss: 0.005824632942676544
step: 420, loss: 0.015045076608657837
step: 430, loss: 0.06326806545257568
step: 440, loss: 0.039327844977378845
step: 450, loss: 0.04277808964252472
step: 460, loss: 0.002417501527816057
step: 470, loss: 0.00523380097001791
step: 480, loss: 0.005545184947550297
step: 490, loss: 0.0085115572437644
step: 500, loss: 0.012818528339266777
step: 510, loss: 0.013328175991773605
step: 520, loss: 0.10348467528820038
step: 530, loss: 0.0015743133844807744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9193245778611632, f1=0.920205319645357,  
best_f1=0.920205319645357  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.05056135728955269  
step: 10, loss: 0.017016034573316574  
step: 20, loss: 0.017345653846859932  
step: 30, loss: 0.0013297229306772351  
step: 40, loss: 0.03182334825396538  
step: 50, loss: 0.01096506416797638  
step: 60, loss: 0.000509593344759196  
step: 70, loss: 0.00016351520025637  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 80, loss: 0.000491084938403219  
step: 90, loss: 0.2109847366809845  
step: 100, loss: 0.0008760818163864315  
step: 110, loss: 0.0005137055413797498  
step: 120, loss: 0.003578073810786009  
step: 130, loss: 0.0017774912994354963  
step: 140, loss: 0.0018245010869577527  
step: 150, loss: 0.0011290879920125008  
step: 160, loss: 0.010594777762889862  
step: 170, loss: 0.06101856008172035  
step: 180, loss: 0.0015688554849475622  
step: 190, loss: 0.0012200273340567946  
step: 200, loss: 0.0015604712534695864  
step: 210, loss: 0.0018101041205227375  
step: 220, loss: 0.00043627378181554377  
step: 230, loss: 0.0012032537488266826  
step: 240, loss: 0.004125046543776989  
step: 250, loss: 0.0014422599924728274  
step: 260, loss: 0.24031758308410645  
step: 270, loss: 0.04074350744485855  
step: 280, loss: 0.03301108628511429  
step: 290, loss: 0.013818379491567612  
step: 300, loss: 0.0027861683629453182  
step: 310, loss: 0.12755335867404938  
step: 320, loss: 0.0031559988856315613  
step: 330, loss: 0.0011115154484286904  
step: 340, loss: 0.001242469996213913  
step: 350, loss: 0.003717300947755575  
step: 360, loss: 0.016041550785303116  
step: 370, loss: 0.000634217110928148  
step: 380, loss: 0.0013407664373517036  
step: 390, loss: 0.020762693136930466  
step: 400, loss: 0.003574127797037363  
step: 410, loss: 0.0005021037650294602
```

```

step: 420, loss: 0.008885720744729042
step: 430, loss: 0.0013019958278164268
step: 440, loss: 0.000783504277933389
step: 450, loss: 0.09911423921585083
step: 460, loss: 0.001477857120335102
step: 470, loss: 0.00039481851854361594
step: 480, loss: 0.0024935954716056585
step: 490, loss: 0.06906936317682266
step: 500, loss: 0.015006794594228268
step: 510, loss: 0.007316289469599724
step: 520, loss: 0.0012543704360723495
step: 530, loss: 0.0042357975617051125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9209783631232361, f1=0.9170593779453347,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020056042820215225
step: 10, loss: 0.0004697235126513988
step: 20, loss: 0.0013082362711429596
step: 30, loss: 0.009697109460830688
step: 40, loss: 0.0006802587886340916
step: 50, loss: 0.0008753340225666761
step: 60, loss: 0.000330878741806373
step: 70, loss: 0.004027936141937971
step: 80, loss: 0.003634918946772814
step: 90, loss: 0.0004851262201555073
step: 100, loss: 0.0003273399779573083
step: 110, loss: 0.0046309130266308784
step: 120, loss: 0.0006055732956156135
step: 130, loss: 0.01781989447772503
step: 140, loss: 0.015198898501694202
step: 150, loss: 0.00027178844902664423
step: 160, loss: 8.273861021734774e-05
step: 170, loss: 0.00035478119389154017
step: 180, loss: 0.00047323619946837425
step: 190, loss: 0.0006418202538043261
step: 200, loss: 0.00020112966012675315
step: 210, loss: 0.00869566760957241
step: 220, loss: 0.0001734383695293218
step: 230, loss: 0.006268067751079798

```

```

step: 240, loss: 0.00018884010205511004
step: 250, loss: 0.0006349214818328619
step: 260, loss: 0.05722050741314888
step: 270, loss: 0.019123541191220284
step: 280, loss: 0.021340463310480118
step: 290, loss: 0.003842275822535157
step: 300, loss: 0.00104148022364825
step: 310, loss: 0.010273996740579605
step: 320, loss: 0.02746405079960823
step: 330, loss: 0.007941427640616894
step: 340, loss: 0.05748395994305611
step: 350, loss: 0.020356550812721252
step: 360, loss: 0.013805553317070007
step: 370, loss: 0.0036165679339319468
step: 380, loss: 0.008422351442277431
step: 390, loss: 0.0085557596758008
step: 400, loss: 0.0007903643418103456
step: 410, loss: 0.00026668625650927424
step: 420, loss: 0.04090361297130585
step: 430, loss: 0.0003667268028948456
step: 440, loss: 0.0010748689528554678
step: 450, loss: 0.021233540028333664
step: 460, loss: 0.0022363835014402866
step: 470, loss: 0.03955201432108879
step: 480, loss: 0.0029889349825680256
step: 490, loss: 0.00024810718605294824
step: 500, loss: 0.004413579124957323
step: 510, loss: 0.0001981916429940611
step: 520, loss: 0.0015643071383237839
step: 530, loss: 0.00675115454941988
/usr/local/lib/python3.7/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=0.9192720485300979, f1=0.9146853146853147,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020155725069344044
step: 10, loss: 0.00300790392793715
step: 20, loss: 0.0004116256604902446
step: 30, loss: 0.005573749076575041
step: 40, loss: 0.00022690740297548473
step: 50, loss: 0.24574220180511475
step: 60, loss: 0.013984373770654202
step: 70, loss: 0.006523373071104288
step: 80, loss: 0.012037412263453007
step: 90, loss: 0.0008677600417286158

```

```
step: 100, loss: 0.023075144737958908
step: 110, loss: 0.0011140126734972
step: 120, loss: 0.0009869749192148447
step: 130, loss: 0.0057289088144898415
step: 140, loss: 0.0007539074867963791
step: 150, loss: 0.0001472686417400837
step: 160, loss: 0.013067755848169327
step: 170, loss: 0.18248780071735382
step: 180, loss: 0.0005174080142751336
step: 190, loss: 0.0006692831520922482
step: 200, loss: 0.0005143347661942244
step: 210, loss: 0.0050813197158277035
step: 220, loss: 0.0017034854972735047
step: 230, loss: 0.00022898902534507215
step: 240, loss: 0.000970510533079505
step: 250, loss: 0.004596128594130278
step: 260, loss: 0.00023586026509292424
step: 270, loss: 9.273570321965963e-05
step: 280, loss: 0.009837353602051735
step: 290, loss: 0.001404995797201991
step: 300, loss: 0.0005228696973063052
step: 310, loss: 0.0006435624090954661
step: 320, loss: 0.015999983996152878
step: 330, loss: 0.00020244828192517161
step: 340, loss: 0.0014733209973201156
step: 350, loss: 0.0021120517048984766
step: 360, loss: 0.002119595417752862
step: 370, loss: 0.00084458984201774
step: 380, loss: 0.10555627942085266
step: 390, loss: 9.547876106807962e-05
step: 400, loss: 8.798747876426205e-05
step: 410, loss: 0.0013234586222097278
step: 420, loss: 0.00044713725219480693
step: 430, loss: 0.004627104848623276
step: 440, loss: 0.003015941707417369
step: 450, loss: 0.0010031358106061816
step: 460, loss: 0.000803449482191354
step: 470, loss: 0.006220887415111065
step: 480, loss: 0.0003566934901755303
step: 490, loss: 0.0004794812703039497
step: 500, loss: 0.0010230988264083862
step: 510, loss: 0.0021739560179412365
step: 520, loss: 0.0015561480540782213
step: 530, loss: 7.879734766902402e-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 7: dev_f1=0.9108635097493036, f1=0.9128630705394191,
```

```
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009858165867626667
step: 10, loss: 0.14209839701652527
step: 20, loss: 0.0071786814369261265
step: 30, loss: 0.005984328221529722
step: 40, loss: 0.001269233413040638
step: 50, loss: 0.00140663783531636
step: 60, loss: 0.0004052727308589965
step: 70, loss: 0.02320103347301483
step: 80, loss: 0.0006055876729078591
step: 90, loss: 0.0026233159005641937
step: 100, loss: 0.0016013465356081724
step: 110, loss: 0.007985701784491539
step: 120, loss: 0.0025336185935884714
step: 130, loss: 0.0009703635587356985
step: 140, loss: 6.584347283933312e-05
step: 150, loss: 0.0032637701369822025
step: 160, loss: 0.0004978961660526693
step: 170, loss: 0.00012337903899606317
step: 180, loss: 0.0001790145324775949
step: 190, loss: 0.0001488883572164923
step: 200, loss: 0.0020519979298114777
step: 210, loss: 0.004660534672439098
step: 220, loss: 0.0025751811917871237
step: 230, loss: 0.001074397237971425
step: 240, loss: 0.00020209378271829337
step: 250, loss: 0.0014749558176845312
step: 260, loss: 0.000365554413292557
step: 270, loss: 7.098019705154002e-05
step: 280, loss: 0.0001705441827652976
step: 290, loss: 0.005458373110741377
step: 300, loss: 0.00011429505684645846
step: 310, loss: 0.03388525918126106
step: 320, loss: 0.00031098470208235085
step: 330, loss: 0.00033398214145563543
step: 340, loss: 0.00039559206925332546
step: 350, loss: 0.011969480663537979
step: 360, loss: 0.00011384493700461462
step: 370, loss: 0.0003484478802420199
step: 380, loss: 0.026913383975625038
step: 390, loss: 0.0003848474589176476
step: 400, loss: 0.009096472524106503
step: 410, loss: 0.005109678953886032
step: 420, loss: 6.609497359022498e-05
step: 430, loss: 0.009435269050300121
```

```

step: 440, loss: 0.00042459016549400985
step: 450, loss: 0.00010086162365041673
step: 460, loss: 0.0016056541353464127
step: 470, loss: 0.0006810522172600031
step: 480, loss: 0.0003468056092970073
step: 490, loss: 0.0009388139587827027
step: 500, loss: 0.004699260927736759
step: 510, loss: 0.00023009271535556763
step: 520, loss: 0.012239744886755943
step: 530, loss: 8.138727571349591e-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=0.9138645785352373, f1=0.911926605504587,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014178490964695811
step: 10, loss: 0.0001418101746821776
step: 20, loss: 0.0038211920764297247
step: 30, loss: 6.30072463536635e-05
step: 40, loss: 0.0004386340733617544
step: 50, loss: 0.00023403532395604998
step: 60, loss: 0.004809191450476646
step: 70, loss: 0.025499753654003143
step: 80, loss: 0.0004657460085581988
step: 90, loss: 0.000791819067671895
step: 100, loss: 0.0018082320457324386
step: 110, loss: 0.000505484058521688
step: 120, loss: 9.543081250740215e-05
step: 130, loss: 8.514085493516177e-05
step: 140, loss: 0.07147352397441864
step: 150, loss: 0.001275557791814208
step: 160, loss: 4.773354157805443e-05
step: 170, loss: 9.039627184392884e-05
step: 180, loss: 9.869793575489894e-05
step: 190, loss: 5.495600271387957e-05
step: 200, loss: 0.00098226813133806
step: 210, loss: 9.717217471916229e-05
step: 220, loss: 0.0014100753469392657
step: 230, loss: 6.698754441458732e-05
step: 240, loss: 0.0004725003964267671
step: 250, loss: 0.007539871148765087
step: 260, loss: 0.00010555781045695767
step: 270, loss: 0.0011122089345008135
step: 280, loss: 4.053337761433795e-05
step: 290, loss: 4.034693483845331e-05

```



```

step: 300, loss: 5.025819336879067e-05
step: 310, loss: 0.00564340827986598
step: 320, loss: 3.657391425804235e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 330, loss: 0.0015116475988179445
step: 340, loss: 0.0005037523224018514
step: 350, loss: 0.00514662079513073
step: 360, loss: 0.006761847995221615
step: 370, loss: 0.00617383886128664
step: 380, loss: 3.663352254079655e-05
step: 390, loss: 0.0001676723186392337
step: 400, loss: 0.0766087993979454
step: 410, loss: 0.0037203533574938774
step: 420, loss: 0.007339778356254101
step: 430, loss: 0.0001614473876543343
step: 440, loss: 0.0004024562949780375
step: 450, loss: 6.13037045695819e-05
step: 460, loss: 7.469340198440477e-05
step: 470, loss: 0.001890111481770873
step: 480, loss: 0.0023882496170699596
step: 490, loss: 5.087663157610223e-05
step: 500, loss: 0.00010534169996390119
step: 510, loss: 0.008507449179887772
step: 520, loss: 3.3075328246923164e-05
step: 530, loss: 5.477154263644479e-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=0.9170506912442397, f1=0.9124087591240876,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014222264871932566
step: 10, loss: 2.4835948352119885e-05
step: 20, loss: 6.621051579713821e-05
step: 30, loss: 0.0016125714173540473
step: 40, loss: 0.0007397933513857424
step: 50, loss: 2.8590939109562896e-05
step: 60, loss: 5.326461177901365e-05
step: 70, loss: 6.722223770339042e-05
step: 80, loss: 0.03803979977965355
step: 90, loss: 0.00020388128177728504
step: 100, loss: 0.0007072544540278614
step: 110, loss: 0.0003171883581671864
step: 120, loss: 0.00016851296823006123
step: 130, loss: 8.194569818442687e-05
step: 140, loss: 0.0014321269700303674

```

```

step: 150, loss: 0.0005745440721511841
step: 160, loss: 0.0003960559261031449
step: 170, loss: 4.007418465334922e-05
step: 180, loss: 0.08342041820287704
step: 190, loss: 0.00012229528510943055
step: 200, loss: 8.287889795610681e-05
step: 210, loss: 3.720311724464409e-05
step: 220, loss: 0.00013617459626402706
step: 230, loss: 0.0007341854507103562
step: 240, loss: 3.998526881332509e-05
step: 250, loss: 0.0004744708421640098
step: 260, loss: 0.00034999795025214553
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 270, loss: 0.0003685594128910452
step: 280, loss: 3.178696351824328e-05
step: 290, loss: 0.00023740940378047526
step: 300, loss: 4.1382347262697294e-05
step: 310, loss: 4.4937158236280084e-05
step: 320, loss: 0.0001875302114058286
step: 330, loss: 0.012863344512879848
step: 340, loss: 5.984463859931566e-05
step: 350, loss: 3.793970245169476e-05
step: 360, loss: 0.0018269215943291783
step: 370, loss: 0.00012856710236519575
step: 380, loss: 0.00010201586701441556
step: 390, loss: 3.3801974495872855e-05
step: 400, loss: 5.385194162954576e-05
step: 410, loss: 6.715460767736658e-05
step: 420, loss: 0.00022115296451374888
step: 430, loss: 0.00011637950228760019
step: 440, loss: 3.339991235407069e-05
step: 450, loss: 9.036662231665105e-05
step: 460, loss: 0.0019251969642937183
step: 470, loss: 2.6024248654721305e-05
step: 480, loss: 0.0006793204811401665
step: 490, loss: 0.03614525869488716
step: 500, loss: 0.0004501568619161844
step: 510, loss: 0.0001060294671333395
step: 520, loss: 0.0008732513524591923
step: 530, loss: 9.557662269799039e-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=0.9137148047229792, f1=0.9107303877366997,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.007021884433925152
step: 10, loss: 0.015197831206023693
step: 20, loss: 8.734241419006139e-05
step: 30, loss: 0.0004791493702214211
step: 40, loss: 0.00023799565678928047
step: 50, loss: 0.0004962047096341848
step: 60, loss: 0.0002229885576525703
step: 70, loss: 0.00020656561537180096
step: 80, loss: 8.749468543101102e-05
step: 90, loss: 0.00015252493903972208
step: 100, loss: 8.021183020900935e-05
step: 110, loss: 6.718483928125352e-05
step: 120, loss: 0.0006578064057976007
step: 130, loss: 0.0003933105617761612
step: 140, loss: 6.585820665350184e-05
step: 150, loss: 0.00026763416826725006
step: 160, loss: 0.002313013654202223
step: 170, loss: 0.0017162031726911664
step: 180, loss: 6.744304846506566e-05
step: 190, loss: 0.000984835671260953
step: 200, loss: 0.007080393843352795
step: 210, loss: 0.00020545402367133647
step: 220, loss: 0.00030302171944640577
step: 230, loss: 0.0003841962607111782
step: 240, loss: 5.0366528739687055e-05
step: 250, loss: 6.477283750427887e-05
step: 260, loss: 3.6722653021570295e-05
step: 270, loss: 0.004139487165957689
step: 280, loss: 0.0007509551360271871
step: 290, loss: 0.0008546830504201353
step: 300, loss: 0.0010197520023211837
step: 310, loss: 0.023273561149835587
step: 320, loss: 0.0028309442568570375
step: 330, loss: 0.0006357788224704564
step: 340, loss: 0.00015943418839015067
step: 350, loss: 0.041885748505592346
step: 360, loss: 0.00014386507973540574
step: 370, loss: 4.723189340438694e-05
step: 380, loss: 2.7369058443582617e-05
step: 390, loss: 0.010449264198541641
step: 400, loss: 3.964728966820985e-05
step: 410, loss: 8.899151725927368e-05
step: 420, loss: 0.0003678349603433162
step: 430, loss: 4.478818300412968e-05
step: 440, loss: 4.2559830035315827e-05
step: 450, loss: 0.00011859540245495737
step: 460, loss: 0.0007811989635229111
step: 470, loss: 0.003590279258787632

```

step: 480, loss: 0.00034442561445757747
step: 490, loss: 0.0003425863978918642
step: 500, loss: 7.83056893851608e-05
step: 510, loss: 3.5083634429611266e-05
step: 520, loss: 8.813269232632592e-05
step: 530, loss: 0.00043548826943151653
/usr/local/lib/python3.7/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=0.9022628791526239, f1=0.9051724137931034,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002445618389174342
step: 10, loss: 0.00017614015087019652
step: 20, loss: 3.4684424463193864e-05
step: 30, loss: 4.2264280637027696e-05
step: 40, loss: 0.0002262645139126107
step: 50, loss: 0.0015790530014783144
step: 60, loss: 5.1576564146671444e-05
step: 70, loss: 0.0024837548844516277
step: 80, loss: 0.0017819881904870272
step: 90, loss: 5.772126678493805e-05
step: 100, loss: 0.0009511944372206926
step: 110, loss: 7.619563257321715e-05
step: 120, loss: 0.00016518679331056774
step: 130, loss: 0.00010694832599256188
step: 140, loss: 5.315564339980483e-05
step: 150, loss: 0.00041183590656146407
step: 160, loss: 0.00019304467423353344
step: 170, loss: 6.333955388981849e-05
step: 180, loss: 0.0007190919714048505
step: 190, loss: 2.0753193894051947e-05
step: 200, loss: 6.809965998400003e-05
step: 210, loss: 0.0008373516029678285
step: 220, loss: 6.286903226282448e-05
step: 230, loss: 0.001046092133037746
step: 240, loss: 0.0007323684985749424
step: 250, loss: 5.275202420307323e-05
step: 260, loss: 0.00019375269766896963
step: 270, loss: 4.561715468298644e-05
step: 280, loss: 2.9142136554582976e-05
step: 290, loss: 0.0007784549379721284
step: 300, loss: 0.00010344007750973105
step: 310, loss: 0.00017383725207764655
step: 320, loss: 6.790012412238866e-05
step: 330, loss: 2.795760883600451e-05

```

```

step: 340, loss: 5.1247377996332943e-05
step: 350, loss: 0.0008337440085597336
step: 360, loss: 0.0006967429653741419
step: 370, loss: 0.0002414939517620951
step: 380, loss: 0.0002719753247220069
step: 390, loss: 8.43660527607426e-05
step: 400, loss: 0.0016031407285481691
step: 410, loss: 5.224382766755298e-05
step: 420, loss: 0.0014073449419811368
step: 430, loss: 0.000221268244786188
step: 440, loss: 0.00038098637014627457
step: 450, loss: 0.0010419946629554033
step: 460, loss: 0.00020335589942988008
step: 470, loss: 0.002195905428379774
step: 480, loss: 0.0004337792342994362
step: 490, loss: 0.00030574615811929107
step: 500, loss: 0.0016898111207410693
step: 510, loss: 5.18947999808006e-05
step: 520, loss: 0.021462157368659973
step: 530, loss: 3.555290822987445e-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=0.9089201877934272, f1=0.9104408352668214,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000387996609788388
step: 10, loss: 0.00016135690384544432
step: 20, loss: 0.0017002663807943463
step: 30, loss: 0.004215329885482788
step: 40, loss: 6.162282807053998e-05
step: 50, loss: 0.00022570327564608306
step: 60, loss: 0.00015890770009718835
step: 70, loss: 3.5790930269286036e-05
step: 80, loss: 7.979964721016586e-05
step: 90, loss: 0.0017744768410921097
step: 100, loss: 2.797224988171365e-05
step: 110, loss: 0.0015128531958907843
step: 120, loss: 0.00010030897101387382
step: 130, loss: 0.002041741507127881
step: 140, loss: 0.00030330076697282493
step: 150, loss: 8.9911169197876e-05
step: 160, loss: 0.00010375397687312216
step: 170, loss: 0.0045013995841145515
step: 180, loss: 3.658084096969105e-05
step: 190, loss: 0.0001352582621620968

```

```

step: 200, loss: 0.00021397742966655642
step: 210, loss: 0.00046499367454089224
step: 220, loss: 4.202029958833009e-05
step: 230, loss: 0.00016955290629994124
step: 240, loss: 5.034443165641278e-05
step: 250, loss: 0.017972884699702263
step: 260, loss: 0.008399806916713715
step: 270, loss: 0.00013651019253302366
step: 280, loss: 0.0023571234196424484
step: 290, loss: 0.0003264553379267454
step: 300, loss: 0.0001331380772171542
step: 310, loss: 9.495862468611449e-05
step: 320, loss: 0.00013296661200001836
step: 330, loss: 0.00012147628876846284
step: 340, loss: 0.0016064730007201433
step: 350, loss: 0.0009342568228021264
step: 360, loss: 3.8047332054702565e-05
step: 370, loss: 7.143663970055059e-05
step: 380, loss: 9.64902137639001e-05
step: 390, loss: 3.081062823184766e-05
step: 400, loss: 4.765574340126477e-05
step: 410, loss: 0.00023511648760177195
step: 420, loss: 0.00012912270904053003
step: 430, loss: 9.878622950054705e-05
step: 440, loss: 0.00017330711125396192
step: 450, loss: 0.002771282335743308
step: 460, loss: 0.0003379913978278637
step: 470, loss: 0.019367249682545662
step: 480, loss: 0.0015549357049167156
step: 490, loss: 7.041186472633854e-05
step: 500, loss: 9.476814011577517e-05
step: 510, loss: 5.348234117263928e-05
step: 520, loss: 2.722732642723713e-05
step: 530, loss: 5.079379479866475e-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=0.9115710253998118, f1=0.9086672879776329,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011987870675511658
step: 10, loss: 0.00013620938989333808
step: 20, loss: 2.542793117754627e-05
step: 30, loss: 8.913968486012891e-05
step: 40, loss: 0.00011295310105197132
step: 50, loss: 7.39464521757327e-05

```

step: 60, loss: 8.235795394284651e-05
step: 70, loss: 3.830187415587716e-05
step: 80, loss: 0.00011023408296750858
step: 90, loss: 4.4924461690243334e-05
step: 100, loss: 0.013625353574752808
step: 110, loss: 3.3403048291802406e-05
step: 120, loss: 2.360620601393748e-05
step: 130, loss: 3.471677337074652e-05
step: 140, loss: 0.00036004045978188515
step: 150, loss: 7.088588608894497e-05
step: 160, loss: 9.633550507714972e-05
step: 170, loss: 0.0006113849813118577
step: 180, loss: 0.0002495306252967566
step: 190, loss: 0.00011460355017334223
step: 200, loss: 5.389963189372793e-05
step: 210, loss: 0.00015546528447885066
step: 220, loss: 3.314997229608707e-05
step: 230, loss: 0.0002510528138373047
step: 240, loss: 3.0665465601487085e-05
step: 250, loss: 0.0003391340433154255
step: 260, loss: 2.267865238536615e-05
step: 270, loss: 0.004913071636110544
step: 280, loss: 2.0000610675197095e-05
step: 290, loss: 6.443206075346097e-05
step: 300, loss: 4.346569039626047e-05
step: 310, loss: 0.00014643203758168966
step: 320, loss: 0.0006591903511434793
step: 330, loss: 2.3997419702936895e-05
step: 340, loss: 7.547704444732517e-05
step: 350, loss: 4.6834662498440593e-05
step: 360, loss: 0.0006743118283338845
step: 370, loss: 9.627713006921113e-05
step: 380, loss: 2.7357171347830445e-05
step: 390, loss: 5.663678894052282e-05
step: 400, loss: 8.096237434074283e-05
step: 410, loss: 0.00022946455283090472
step: 420, loss: 2.5282828573836014e-05
step: 430, loss: 4.751324013341218e-05
step: 440, loss: 2.604167821118608e-05
step: 450, loss: 5.7619578001322225e-05
step: 460, loss: 0.0003935419663321227
step: 470, loss: 2.125585888279602e-05
step: 480, loss: 0.00016208119632210582
step: 490, loss: 3.1226943974616006e-05
step: 500, loss: 4.226157034281641e-05
step: 510, loss: 3.754876888706349e-05
step: 520, loss: 6.227345875231549e-05
step: 530, loss: 0.0002344934327993542

```

/usr/local/lib/python3.7/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=0.9112595419847328, f1=0.9096071935636536,
best_f1=0.9170593779453347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014918872620910406
step: 10, loss: 4.701961006503552e-05
step: 20, loss: 3.46131892001722e-05
step: 30, loss: 6.674609903711826e-05
step: 40, loss: 9.624251106288284e-05
step: 50, loss: 3.6106597690377384e-05
step: 60, loss: 2.07158245757455e-05
step: 70, loss: 3.471508534858003e-05
step: 80, loss: 9.56917938310653e-05
step: 90, loss: 3.7691777833970264e-05
step: 100, loss: 2.055948607448954e-05
step: 110, loss: 0.0016086179530248046
step: 120, loss: 0.0034855862613767385
step: 130, loss: 3.0938688723836094e-05
step: 140, loss: 2.1710433429689147e-05
step: 150, loss: 5.813976167701185e-05
step: 160, loss: 3.141757406410761e-05
step: 170, loss: 1.983309630304575e-05
step: 180, loss: 2.8663947887253016e-05
step: 190, loss: 7.145322888391092e-05
step: 200, loss: 7.528747664764524e-05
step: 210, loss: 0.000218788962229155
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 220, loss: 1.706138027657289e-05
step: 230, loss: 2.2716130843036808e-05
step: 240, loss: 2.6802830689121038e-05
step: 250, loss: 4.319579602451995e-05
step: 260, loss: 5.418344881036319e-05
step: 270, loss: 0.0011191919911652803
step: 280, loss: 0.00011424126569181681
step: 290, loss: 0.0010923452209681273
step: 300, loss: 0.022028597071766853
step: 310, loss: 5.03898918395862e-05
step: 320, loss: 9.265608969144523e-05
step: 330, loss: 0.005254808813333511
step: 340, loss: 0.00024008192121982574
step: 350, loss: 6.240588118089363e-05
step: 360, loss: 0.0002087000320898369
step: 370, loss: 2.1330604795366526e-05
step: 380, loss: 2.2492567950394005e-05

```



```

step: 390, loss: 2.519300687708892e-05
step: 400, loss: 8.490897744195536e-05
step: 410, loss: 0.0017156677786260843
step: 420, loss: 7.445011578965932e-05
step: 430, loss: 2.7372396289138123e-05
step: 440, loss: 7.491902215406299e-05
step: 450, loss: 0.0010010723490267992
step: 460, loss: 1.9508846889948472e-05
step: 470, loss: 2.9982189516886137e-05
step: 480, loss: 3.1512339774053544e-05
step: 490, loss: 4.701817306340672e-05
step: 500, loss: 3.0665203667012975e-05
step: 510, loss: 0.00018734941841103137
step: 520, loss: 3.961298716603778e-05
step: 530, loss: 0.0003376147651579231
/usr/local/lib/python3.7/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=0.9131255901794145, f1=0.9127579737335835,
best_f1=0.9170593779453347
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:21, 264.10it/s]
load_f1 = 0.9169000933706816
real_f1 = 0.91728624535316
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:16, 271.19it/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 distilbert \
  --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 distilbert \

```

```
--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: 0.8785133361816406

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=0.3448275862068965, f1=0.1, best_f1=0.1
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3800339698791504
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4324324324324324, f1=0.16, best_f1=0.16
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30668431520462036
/usr/local/lib/python3.7/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=0.4324324324324324, f1=0.29629629629629634, best_f1=0.16
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.337671160697937
/usr/local/lib/python3.7/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=0.4210526315789474, f1=0.43137254901960786, best_f1=0.16
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2653988003730774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45161290322580644, f1=0.25, best_f1=0.25
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29468101263046265
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5, f1=0.1951219512195122, best_f1=0.1951219512195122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19409382343292236
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.588235294117647, f1=0.2857142857142857,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3486303985118866
/usr/local/lib/python3.7/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=0.5641025641025641, f1=0.24242424242424243,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14871802926063538
/usr/local/lib/python3.7/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=0.5806451612903226, f1=0.2857142857142857,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2346552163362503
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.3243243243243243,
best_f1=0.2857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2640063762664795
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6111111111111112, f1=0.26666666666666666,
best_f1=0.26666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2511509656906128
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6206896551724138, 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: 0.2001207172870636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.631578947368421, f1=0.3225806451612903,
best_f1=0.3225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19264012575149536
/usr/local/lib/python3.7/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=0.631578947368421, f1=0.3225806451612903,
best_f1=0.3225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2764603793621063
/usr/local/lib/python3.7/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=0.631578947368421, f1=0.3225806451612903,
best_f1=0.3225806451612903
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, 130266.78it/s]
/usr/local/lib/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 = 0.5517241379310344
real_f1 = 0.5142857142857143
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:16, 271.70it/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 distilbert \
--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 distilbert \
--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: 0.8007489442825317
step: 10, loss: 0.4737977087497711
step: 20, loss: 0.5803298354148865
step: 30, loss: 0.48605868220329285
step: 40, loss: 0.35076776146888733
step: 50, loss: 0.234954372048378
step: 60, loss: 0.10785814374685287
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.22592391073703766
step: 80, loss: 0.23994246125221252
step: 90, loss: 0.039541251957416534
step: 100, loss: 0.05835006758570671
step: 110, loss: 0.05835181474685669
step: 120, loss: 0.06861317902803421
step: 130, loss: 0.05485185608267784
step: 140, loss: 0.020923588424921036
step: 150, loss: 0.04919992387294769
step: 160, loss: 0.14704249799251556
step: 170, loss: 0.03804415836930275
step: 180, loss: 0.009651610627770424
step: 190, loss: 0.01488885935395956
step: 200, loss: 0.027795037254691124
step: 210, loss: 0.05494567006826401

```

```

step: 220, loss: 0.05423862114548683
step: 230, loss: 0.1172514259815216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9765363128491621, f1=0.9671574178935448,
best_f1=0.9671574178935448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2153501808643341
step: 10, loss: 0.008316894061863422
step: 20, loss: 0.004458664916455746
step: 30, loss: 0.006708574015647173
step: 40, loss: 0.07099101692438126
step: 50, loss: 0.002931684022769332
step: 60, loss: 0.00254421541467309
step: 70, loss: 0.02823510207235813
step: 80, loss: 0.004144573118537664
step: 90, loss: 0.0888315960764885
step: 100, loss: 0.07436912506818771
step: 110, loss: 0.011328148655593395
step: 120, loss: 0.006331630516797304
step: 130, loss: 0.0076712691225111485
step: 140, loss: 0.045297835022211075
step: 150, loss: 0.06195220723748207
step: 160, loss: 0.006996935699135065
step: 170, loss: 0.039021946489810944
step: 180, loss: 0.00216536782681942
step: 190, loss: 0.059107907116413116
step: 200, loss: 0.030161170288920403
step: 210, loss: 0.09791012853384018
step: 220, loss: 0.025455772876739502
step: 230, loss: 0.02805979922413826
/usr/local/lib/python3.7/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=0.9736540664375716, f1=0.9643268124280783,
best_f1=0.9671574178935448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13621944189071655
step: 10, loss: 0.013106687925755978

```

```

step: 20, loss: 0.00226185517385602
step: 30, loss: 0.033438604325056076
step: 40, loss: 0.03721105307340622
step: 50, loss: 0.0033449814654886723
step: 60, loss: 0.009733790531754494
step: 70, loss: 0.0006341426051221788
step: 80, loss: 0.08355525881052017
step: 90, loss: 0.022663172334432602
step: 100, loss: 0.0054208626970648766
step: 110, loss: 0.0019133791793137789
step: 120, loss: 0.021303433924913406
step: 130, loss: 0.0031085433438420296
step: 140, loss: 0.009185511618852615
step: 150, loss: 0.005061538890004158
step: 160, loss: 0.0038684653118252754
step: 170, loss: 0.03182784467935562
step: 180, loss: 0.003213623072952032
step: 190, loss: 0.0030892444774508476
step: 200, loss: 0.020655306056141853
step: 210, loss: 0.04104997590184212
step: 220, loss: 0.005655257496982813
step: 230, loss: 0.05562485009431839
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.978675645342312, f1=0.9742441209406495,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001865419908426702
step: 10, loss: 0.001306062564253807
step: 20, loss: 0.0013305413303896785
step: 30, loss: 0.00045956080430187285
step: 40, loss: 0.0038927202112972736
step: 50, loss: 0.000750472885556519
step: 60, loss: 0.09100836515426636
step: 70, loss: 0.08822232484817505
step: 80, loss: 0.025155460461974144
step: 90, loss: 0.006636433769017458
step: 100, loss: 0.005682711489498615
step: 110, loss: 0.0265577994287014
step: 120, loss: 0.001501427381299436
step: 130, loss: 0.0020204901229590178

```

```

step: 140, loss: 0.0017670791130512953
step: 150, loss: 0.0016138778300955892
step: 160, loss: 0.0005854701739735901
step: 170, loss: 0.07127805799245834
step: 180, loss: 0.04360348731279373
step: 190, loss: 0.004073670133948326
step: 200, loss: 0.0017863715765997767
step: 210, loss: 0.00752261420711875
step: 220, loss: 0.0003671740705613047
step: 230, loss: 0.0003481132735032588
/usr/local/lib/python3.7/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=0.9784824462061155, f1=0.9748283752860413,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006736059440299869
step: 10, loss: 0.0003786178713198751
step: 20, loss: 0.00029501592507585883
step: 30, loss: 0.00011600811558309942
step: 40, loss: 0.0003209798887837678
step: 50, loss: 0.00023939179664012045
step: 60, loss: 0.00020216470875311643
step: 70, loss: 0.00028275049407966435
step: 80, loss: 0.00068571517476812
step: 90, loss: 0.0010501914657652378
step: 100, loss: 0.004686454311013222
step: 110, loss: 0.0026166606694459915
step: 120, loss: 0.03444591909646988
step: 130, loss: 0.0024137741420418024
step: 140, loss: 0.0005132619407959282
step: 150, loss: 0.00024309901345986873
step: 160, loss: 0.0010623389389365911
step: 170, loss: 0.003437599865719676
step: 180, loss: 0.004770966712385416
step: 190, loss: 0.00016739917919039726
step: 200, loss: 0.0015917143318802118
step: 210, loss: 0.1345018893480301
step: 220, loss: 0.0007886969833634794
step: 230, loss: 0.030364196747541428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.

```

```

    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.9797752808988766, f1=0.9718785151856018,
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: 0.0022105295211076736
step: 10, loss: 0.0062681157141923904
step: 20, loss: 0.0017748407553881407
step: 30, loss: 0.006237556226551533
step: 40, loss: 0.0072584995068609715
step: 50, loss: 0.10130187124013901
step: 60, loss: 0.034442752599716187
step: 70, loss: 0.00041956541826948524
step: 80, loss: 0.0020057931542396545
step: 90, loss: 0.0031969358678907156
step: 100, loss: 0.00033214877475984395
step: 110, loss: 0.02004009671509266
step: 120, loss: 0.00047242295113392174
step: 130, loss: 0.004351624753326178
step: 140, loss: 0.0017277925508096814
step: 150, loss: 0.010769371874630451
step: 160, loss: 0.0007494958699680865
step: 170, loss: 0.00046439358266070485
step: 180, loss: 0.0008715035510249436
step: 190, loss: 0.01262866985052824
step: 200, loss: 0.1660814732313156
step: 210, loss: 0.0011949874460697174
step: 220, loss: 0.005959527567028999
step: 230, loss: 0.0004022317298222333
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9764309764309763,
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: 0.008423404768109322
step: 10, loss: 0.004958954639732838
step: 20, loss: 0.0020570429041981697
step: 30, loss: 0.0018368541495874524
step: 40, loss: 0.003725294955074787
step: 50, loss: 0.0015715783229097724
step: 60, loss: 0.0002740110212471336
step: 70, loss: 0.16382993757724762
step: 80, loss: 0.0010249658953398466
step: 90, loss: 0.0004974159528501332

```

```

step: 100, loss: 0.03406267613172531
step: 110, loss: 0.0015267586568370461
step: 120, loss: 0.007444108370691538
step: 130, loss: 0.000836517137940973
step: 140, loss: 0.0007991805905476213
step: 150, loss: 0.00032352155540138483
step: 160, loss: 0.0004516916524153203
step: 170, loss: 0.000158595823450014
step: 180, loss: 0.00043497487786225975
step: 190, loss: 0.0002281149209011346
step: 200, loss: 0.035833410918712616
step: 210, loss: 0.0022585480473935604
step: 220, loss: 0.002105223946273327
step: 230, loss: 0.007671324536204338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.9820224719101124, f1=0.9711111111111111,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01115085743367672
step: 10, loss: 0.1393541693687439
step: 20, loss: 0.00014039562665857375
step: 30, loss: 0.00041067160782404244
step: 40, loss: 0.0003167613467667252
step: 50, loss: 0.002466216217726469
step: 60, loss: 0.0006142751080915332
step: 70, loss: 0.00020728024537675083
step: 80, loss: 0.00017902614490594715
step: 90, loss: 0.0029640933498740196
step: 100, loss: 0.00040651808376424015
step: 110, loss: 0.0045514535158872604
step: 120, loss: 0.0011356206377968192
step: 130, loss: 0.027839139103889465
step: 140, loss: 0.00028066514641977847
step: 150, loss: 0.042847055941820145
step: 160, loss: 0.00042971523362211883
step: 170, loss: 0.0013434465508908033
step: 180, loss: 0.005754278972744942
step: 190, loss: 0.0008292482234537601
step: 200, loss: 0.0047090318985283375
step: 210, loss: 0.0015405337326228619

```

```

step: 220, loss: 0.0002957966935355216
step: 230, loss: 0.0004723906167782843
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.9752252252252253,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009933612309396267
step: 10, loss: 0.00014437257777899504
step: 20, loss: 0.0007775566773489118
step: 30, loss: 0.0003086476936005056
step: 40, loss: 0.0029489025473594666
step: 50, loss: 0.00045421833056025207
step: 60, loss: 0.0003089032252319157
step: 70, loss: 0.004445691127330065
step: 80, loss: 0.02225509285926819
step: 90, loss: 0.00031019875314086676
step: 100, loss: 9.871240763459355e-05
step: 110, loss: 0.00019520540081430227
step: 120, loss: 0.003473482094705105
step: 130, loss: 0.00011709459067787975
step: 140, loss: 0.00013674328511115164
step: 150, loss: 0.00012324145063757896
step: 160, loss: 0.004798418842256069
step: 170, loss: 0.00018685155373532325
step: 180, loss: 0.00019610834715422243
step: 190, loss: 0.11727027595043182
step: 200, loss: 6.86302300891839e-05
step: 210, loss: 0.00017723256314639002
step: 220, loss: 0.042232364416122437
step: 230, loss: 0.0014464829582720995
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9763779527559054,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004935259930789471
step: 10, loss: 0.013101967982947826
step: 20, loss: 0.00023537893139291555
step: 30, loss: 0.0009564297506585717
step: 40, loss: 0.001689059310592711
step: 50, loss: 0.03413930535316467

```

```

step: 60, loss: 0.0014849392464384437
step: 70, loss: 0.0007794175180606544
step: 80, loss: 0.024385446682572365
step: 90, loss: 0.00019473102292977273
step: 100, loss: 0.0001009888292173855
step: 110, loss: 0.00016915977175813168
step: 120, loss: 0.0011610400397330523
step: 130, loss: 0.0010129811707884073
step: 140, loss: 0.1203199103474617
step: 150, loss: 0.0036309172865003347
step: 160, loss: 0.0010907058604061604
step: 170, loss: 0.0056839026510715485
step: 180, loss: 0.023406272754073143
step: 190, loss: 0.0030132054816931486
step: 200, loss: 0.0001858818723121658
step: 210, loss: 0.00022446805087383837
step: 220, loss: 0.0005787370027974248
step: 230, loss: 0.0014904610579833388
/usr/local/lib/python3.7/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=0.9785310734463276, f1=0.9797297297297298,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015027073095552623
step: 10, loss: 0.001063288189470768
step: 20, loss: 7.066733087413013e-05
step: 30, loss: 0.0015465831384062767
step: 40, loss: 0.0002605953486636281
step: 50, loss: 0.0012730715097859502
step: 60, loss: 0.0006537075969390571
step: 70, loss: 0.003950308542698622
step: 80, loss: 0.0011852695606648922
step: 90, loss: 0.0005217841826379299
step: 100, loss: 0.00023692124523222446
step: 110, loss: 9.687769488664344e-05
step: 120, loss: 0.0006492419634014368
step: 130, loss: 0.00019133748719468713
step: 140, loss: 0.0005695644649676979
step: 150, loss: 0.00024226475215982646
step: 160, loss: 0.000158956871018745
step: 170, loss: 0.00026757532032206655
step: 180, loss: 0.005133738741278648
step: 190, loss: 0.01808527298271656
step: 200, loss: 0.0002344495296711102
step: 210, loss: 0.0002351493894821033

```



```

step: 220, loss: 9.231899457518011e-05
step: 230, loss: 0.00018470644135959446
/usr/local/lib/python3.7/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=0.9774266365688488, f1=0.9786276715410572,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029481062665581703
step: 10, loss: 0.0006545427604578435
step: 20, loss: 7.276350515894592e-05
step: 30, loss: 6.055002813809551e-05
step: 40, loss: 0.00014358651242218912
step: 50, loss: 0.000160362251335755
step: 60, loss: 0.00048569196951575577
step: 70, loss: 0.00014963641297072172
step: 80, loss: 0.0004716615076176822
step: 90, loss: 5.709569450118579e-05
step: 100, loss: 0.0009749839082360268
step: 110, loss: 0.0002811890735756606
step: 120, loss: 0.00014581058348994702
step: 130, loss: 5.15103674842976e-05
step: 140, loss: 0.00010094928438775241
step: 150, loss: 0.00016140718071255833
step: 160, loss: 0.0004929230199195445
step: 170, loss: 0.00037449569208547473
step: 180, loss: 4.2146333726122975e-05
step: 190, loss: 0.00033295058528892696
step: 200, loss: 0.00044455454917624593
step: 210, loss: 5.6207187299150974e-05
step: 220, loss: 0.0005928409518674016
step: 230, loss: 0.0003720089152920991
/usr/local/lib/python3.7/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=0.9796839729119639, f1=0.9796380090497738,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001437869854271412
step: 10, loss: 0.00048273502034135163
step: 20, loss: 0.0001319724105997011
step: 30, loss: 0.0021797390654683113
step: 40, loss: 0.0004893467994406819
step: 50, loss: 0.00011390420695533976

```

```

step: 60, loss: 0.0001457061734981835
step: 70, loss: 5.356655674404465e-05
step: 80, loss: 0.00015041018195915967
step: 90, loss: 6.779547402402386e-05
step: 100, loss: 0.000570435484405607
step: 110, loss: 5.923537537455559e-05
step: 120, loss: 7.9638535680715e-05
step: 130, loss: 0.0002820380323100835
step: 140, loss: 0.000168363461853005
step: 150, loss: 0.04306923970580101
step: 160, loss: 4.119292862014845e-05
step: 170, loss: 8.031151810428128e-05
step: 180, loss: 0.0006307607982307673
step: 190, loss: 7.933594315545633e-05
step: 200, loss: 0.00013436219887807965
step: 210, loss: 3.939389353035949e-05
step: 220, loss: 6.795301305828616e-05
step: 230, loss: 0.00012298914953134954
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9786276715410572,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013607216533273458
step: 10, loss: 2.6944588171318173e-05
step: 20, loss: 0.001396730076521635
step: 30, loss: 8.016591891646385e-05
step: 40, loss: 3.461453161435202e-05
step: 50, loss: 7.295189425349236e-05
step: 60, loss: 0.00013069271517451853
step: 70, loss: 0.0004189085739199072
step: 80, loss: 0.0005398854846134782
step: 90, loss: 0.00020986890012864023
step: 100, loss: 0.0006985518848523498
step: 110, loss: 0.028381384909152985
step: 120, loss: 0.0013533650198951364
step: 130, loss: 8.636020356789231e-05
step: 140, loss: 9.132587001658976e-05
step: 150, loss: 9.951623360393569e-05
step: 160, loss: 7.570378511445597e-05
step: 170, loss: 4.346030254964717e-05
step: 180, loss: 0.00014644459588453174
step: 190, loss: 4.2397528886795044e-05
step: 200, loss: 4.109191650059074e-05
step: 210, loss: 9.219835192197934e-05

```

```

step: 220, loss: 4.3468498915899545e-05
step: 230, loss: 7.528773130616173e-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=0.9785794813979707, f1=0.9774774774774775,
best_f1=0.9711111111111111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.738412149483338e-05
step: 10, loss: 0.00013155100168660283
step: 20, loss: 3.963177732657641e-05
step: 30, loss: 7.488441042369232e-05
step: 40, loss: 3.837297481368296e-05
step: 50, loss: 3.911817475454882e-05
step: 60, loss: 0.00012067847274011001
step: 70, loss: 6.248097633942962e-05
step: 80, loss: 6.932279211468995e-05
step: 90, loss: 3.388067489140667e-05
step: 100, loss: 0.00010418020974611863
step: 110, loss: 0.0001003327124635689
step: 120, loss: 3.9740138163324445e-05
step: 130, loss: 7.650915358681232e-05
step: 140, loss: 4.396416625240818e-05
step: 150, loss: 9.226116526406258e-05
step: 160, loss: 4.035393430967815e-05
step: 170, loss: 3.180187559337355e-05
step: 180, loss: 0.0007685926975682378
step: 190, loss: 0.0001295939728152007
step: 200, loss: 5.8229117712471634e-05
step: 210, loss: 0.005992108955979347
step: 220, loss: 0.00022965189418755472
step: 230, loss: 0.00022315986279863864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9785794813979707, f1=0.978675645342312,
best_f1=0.9711111111111111
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 [00:10, 245.06it/s]
load_f1 = 0.9820224719101124
real_f1 = 0.9797297297297298
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 267.82it/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 distilbert \
  --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 distilbert \
  --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: 0.7952263355255127

step: 10, loss: 0.41010981798171997

step: 20, loss: 0.4974774122238159

step: 30, loss: 0.4622398614883423

step: 40, loss: 0.3904871940612793

step: 50, loss: 0.2899174988269806

step: 60, loss: 0.32926443219184875

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

step: 70, loss: 0.13476620614528656

step: 80, loss: 0.09165332466363907

step: 90, loss: 0.1323481798171997

step: 100, loss: 0.3068419098854065

step: 110, loss: 0.07460439950227737

step: 120, loss: 0.06351007521152496

step: 130, loss: 0.029723379760980606

step: 140, loss: 0.24088793992996216

step: 150, loss: 0.01960458979010582

step: 160, loss: 0.0998014584183693

step: 170, loss: 0.32628920674324036

step: 180, loss: 0.11867254227399826

step: 190, loss: 0.07794106006622314

step: 200, loss: 0.18586184084415436

step: 210, loss: 0.116609588265419

step: 220, loss: 0.14950378239154816

step: 230, loss: 0.10450620204210281

step: 240, loss: 0.14770448207855225

step: 250, loss: 0.15632012486457825

step: 260, loss: 0.023733438923954964

step: 270, loss: 0.04433068260550499

step: 280, loss: 0.2664106488227844

step: 290, loss: 0.048638928681612015

step: 300, loss: 0.07752270996570587

step: 310, loss: 0.10659024864435196

step: 320, loss: 0.046191614121198654

step: 330, loss: 0.1690634787082672

step: 340, loss: 0.15362559258937836

step: 350, loss: 0.17141713201999664

step: 360, loss: 0.1045951098203659

step: 370, loss: 0.21538464725017548

step: 380, loss: 0.2361607700586319

```

step: 390, loss: 0.10865215212106705
step: 400, loss: 0.05767349526286125
step: 410, loss: 0.05557820200920105
step: 420, loss: 0.03326458856463432
step: 430, loss: 0.06909376382827759
step: 440, loss: 0.13926811516284943
step: 450, loss: 0.014197330921888351
step: 460, loss: 0.05266539007425308
step: 470, loss: 0.22746792435646057
step: 480, loss: 0.27871599793434143
step: 490, loss: 0.061469241976737976
step: 500, loss: 0.045078691095113754
step: 510, loss: 0.0777871236205101
step: 520, loss: 0.1610596776008606
step: 530, loss: 0.17538633942604065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9040472942246477, f1=0.9008189262966332,
best_f1=0.9008189262966332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13639362156391144
step: 10, loss: 0.17072944343090057
step: 20, loss: 0.19683735072612762
step: 30, loss: 0.07915789633989334
step: 40, loss: 0.0536065436899662
step: 50, loss: 0.1262415647506714
step: 60, loss: 0.17397986352443695
step: 70, loss: 0.20073369145393372
step: 80, loss: 0.024274766445159912
step: 90, loss: 0.08098851889371872
step: 100, loss: 0.27284181118011475
step: 110, loss: 0.06685560941696167
step: 120, loss: 0.1306951493024826
step: 130, loss: 0.12562039494514465
step: 140, loss: 0.03454515337944031
step: 150, loss: 0.08002053946256638
step: 160, loss: 0.042961955070495605
step: 170, loss: 0.08876460045576096
step: 180, loss: 0.011200843378901482
step: 190, loss: 0.1385519951581955
step: 200, loss: 0.024805104359984398

```

```

step: 210, loss: 0.008280440233647823
step: 220, loss: 0.12976297736167908
step: 230, loss: 0.05918009579181671
step: 240, loss: 0.1964590847492218
step: 250, loss: 0.06722728908061981
step: 260, loss: 0.041498471051454544
step: 270, loss: 0.15332214534282684
step: 280, loss: 0.16003857553005219
step: 290, loss: 0.16054774820804596
step: 300, loss: 0.03676855191588402
step: 310, loss: 0.0474262572824955
step: 320, loss: 0.0960659384727478
step: 330, loss: 0.039951473474502563
step: 340, loss: 0.011987724341452122
step: 350, loss: 0.04961752891540527
step: 360, loss: 0.10545308142900467
step: 370, loss: 0.003939946182072163
step: 380, loss: 0.09755164384841919
step: 390, loss: 0.04693980515003204
step: 400, loss: 0.1506272554397583
step: 410, loss: 0.03487132117152214
step: 420, loss: 0.06875636428594589
step: 430, loss: 0.040529001504182816
step: 440, loss: 0.03146407753229141
step: 450, loss: 0.02051376923918724
step: 460, loss: 0.161459818482399
step: 470, loss: 0.08192593604326248
step: 480, loss: 0.1419304758310318
step: 490, loss: 0.06967896968126297
step: 500, loss: 0.029286161065101624
step: 510, loss: 0.040824562311172485
step: 520, loss: 0.025472210720181465
step: 530, loss: 0.10996222496032715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.905642484589853, f1=0.9071969696969696,
best_f1=0.9071969696969696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10545818507671356
step: 10, loss: 0.17905427515506744
step: 20, loss: 0.17209921777248383

```


step: 30, loss: 0.2123388946056366
step: 40, loss: 0.06218170002102852
step: 50, loss: 0.012952693738043308
step: 60, loss: 0.015975112095475197
step: 70, loss: 0.043676964938640594
step: 80, loss: 0.10348618030548096
step: 90, loss: 0.04695401340723038
step: 100, loss: 0.09761510044336319
step: 110, loss: 0.03501126542687416
step: 120, loss: 0.15073572099208832
step: 130, loss: 0.07075721025466919
step: 140, loss: 0.04332777112722397
step: 150, loss: 0.0038260030560195446
step: 160, loss: 0.04288007691502571
step: 170, loss: 0.06709518283605576
step: 180, loss: 0.08569198101758957
step: 190, loss: 0.022592518478631973
step: 200, loss: 0.02310085855424404
step: 210, loss: 0.04330037534236908
step: 220, loss: 0.08933259546756744
step: 230, loss: 0.009109249338507652
step: 240, loss: 0.010742666199803352
step: 250, loss: 0.014530820772051811
step: 260, loss: 0.0022757030092179775
step: 270, loss: 0.0052205040119588375
step: 280, loss: 0.038357608020305634
step: 290, loss: 0.04970667511224747
step: 300, loss: 0.08044622838497162
step: 310, loss: 0.06399159878492355
step: 320, loss: 0.1661047786474228
step: 330, loss: 0.013456516899168491
step: 340, loss: 0.010878115892410278
step: 350, loss: 0.026310289278626442
step: 360, loss: 0.013416709378361702
step: 370, loss: 0.0065901437774300575
step: 380, loss: 0.026403209194540977
step: 390, loss: 0.009751083329319954
step: 400, loss: 0.0857384130358696
step: 410, loss: 0.013397280126810074
step: 420, loss: 0.027954667806625366
step: 430, loss: 0.024904994294047356
step: 440, loss: 0.17752233147621155
step: 450, loss: 0.05587742105126381
step: 460, loss: 0.2170974165201187
step: 470, loss: 0.011879881843924522
step: 480, loss: 0.014408765360713005
step: 490, loss: 0.08897310495376587
step: 500, loss: 0.26726382970809937

```

step: 510, loss: 0.006514896173030138
step: 520, loss: 0.0465669147670269
step: 530, loss: 0.15475907921791077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.912901723334886, f1=0.9130841121495327,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006990138906985521
step: 10, loss: 0.07109255343675613
step: 20, loss: 0.10618709772825241
step: 30, loss: 0.04794232174754143
step: 40, loss: 0.0054271044209599495
step: 50, loss: 0.03659603372216225
step: 60, loss: 0.003942370880395174
step: 70, loss: 0.0018674257444217801
step: 80, loss: 0.011533279903233051
step: 90, loss: 0.013758835382759571
step: 100, loss: 0.008023792877793312
step: 110, loss: 0.05574401468038559
step: 120, loss: 0.004774733912199736
step: 130, loss: 0.08754213154315948
step: 140, loss: 0.026300493627786636
step: 150, loss: 0.0037422163877636194
step: 160, loss: 0.008804881013929844
step: 170, loss: 0.004792890045791864
step: 180, loss: 0.03852927312254906
step: 190, loss: 0.031289223581552505
step: 200, loss: 0.0047286939807236195
step: 210, loss: 0.18883392214775085
step: 220, loss: 0.002659645862877369
step: 230, loss: 0.27575942873954773
step: 240, loss: 0.17147748172283173
step: 250, loss: 0.028176169842481613
step: 260, loss: 0.0845838263630867
step: 270, loss: 0.06226395070552826
step: 280, loss: 0.0038566545117646456
step: 290, loss: 0.04218866303563118
step: 300, loss: 0.0018719563959166408
step: 310, loss: 0.001706511597149074
step: 320, loss: 0.027379663661122322

```

```

step: 330, loss: 0.2475501447916031
step: 340, loss: 0.1542791724205017
step: 350, loss: 0.009873994626104832
step: 360, loss: 0.01060103252530098
step: 370, loss: 0.00535841379314661
step: 380, loss: 0.004461426287889481
step: 390, loss: 0.037696611136198044
step: 400, loss: 0.03306897357106209
step: 410, loss: 0.05907789245247841
step: 420, loss: 0.005134066101163626
step: 430, loss: 0.052348099648952484
step: 440, loss: 0.13453266024589539
step: 450, loss: 0.15067848563194275
step: 460, loss: 0.07883643358945847
step: 470, loss: 0.0025188291911035776
step: 480, loss: 0.010651652701199055
step: 490, loss: 0.0776175931096077
step: 500, loss: 0.06761123985052109
step: 510, loss: 0.04238073527812958
step: 520, loss: 0.015465510077774525
step: 530, loss: 0.0006069181254133582
/usr/local/lib/python3.7/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=0.9044585987261148, f1=0.9074410163339383,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04569011926651001
step: 10, loss: 0.005994804203510284
step: 20, loss: 0.006873711012303829
step: 30, loss: 0.009259842336177826
step: 40, loss: 0.030485549941658974
step: 50, loss: 0.007677775342017412
step: 60, loss: 0.0007639203104190528
step: 70, loss: 0.048086583614349365
step: 80, loss: 0.004350699484348297
step: 90, loss: 0.011951243504881859
step: 100, loss: 0.0012399133993312716
step: 110, loss: 0.001584571786224842
step: 120, loss: 0.001572661567479372
step: 130, loss: 0.007736958563327789
step: 140, loss: 0.0011682562762871385
step: 150, loss: 0.00417561549693346
step: 160, loss: 0.05769338086247444
step: 170, loss: 0.013247926719486713
step: 180, loss: 0.008526827208697796

```

```
step: 190, loss: 0.002491182880476117
step: 200, loss: 0.007223836146295071
step: 210, loss: 0.0015880396822467446
step: 220, loss: 0.0008666340727359056
step: 230, loss: 0.0015947232022881508
step: 240, loss: 0.04412836208939552
step: 250, loss: 0.0008932057535275817
step: 260, loss: 0.006584086921066046
step: 270, loss: 0.006984703242778778
step: 280, loss: 0.052168913185596466
step: 290, loss: 0.0012536385329440236
step: 300, loss: 0.16215959191322327
step: 310, loss: 0.0012744696578010917
step: 320, loss: 0.015792973339557648
step: 330, loss: 0.0633661076426506
step: 340, loss: 0.062183678150177
step: 350, loss: 0.02621670812368393
step: 360, loss: 0.11903229355812073
step: 370, loss: 0.024854062125086784
step: 380, loss: 0.0712440088391304
step: 390, loss: 0.048710208386182785
step: 400, loss: 0.011502775363624096
step: 410, loss: 0.006072867661714554
step: 420, loss: 0.011123734526336193
step: 430, loss: 0.024433225393295288
step: 440, loss: 0.0038490821607410908
step: 450, loss: 0.03695942834019661
step: 460, loss: 0.0011229009833186865
step: 470, loss: 0.016994083300232887
step: 480, loss: 0.038963623344898224
step: 490, loss: 0.010694852098822594
step: 500, loss: 0.05967342108488083
step: 510, loss: 0.005744906608015299
step: 520, loss: 0.0012929628137499094
step: 530, loss: 0.04079725965857506
/usr/local/lib/python3.7/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=0.906679298910469, f1=0.9054820415879017,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019326671957969666
step: 10, loss: 0.035543158650398254
step: 20, loss: 0.011407882906496525
step: 30, loss: 0.0015729869483038783
step: 40, loss: 0.0009770262986421585
```

step: 50, loss: 0.003401900641620159
step: 60, loss: 0.004269298631697893
step: 70, loss: 0.048501741141080856
step: 80, loss: 0.0006358937243930995
step: 90, loss: 0.001175197772681713
step: 100, loss: 0.02806822583079338
step: 110, loss: 0.0041935620829463005
step: 120, loss: 0.0030479049310088158
step: 130, loss: 0.0025383243337273598
step: 140, loss: 0.006153434049338102
step: 150, loss: 0.006988204084336758
step: 160, loss: 0.06527936458587646
step: 170, loss: 0.0005453407065942883
step: 180, loss: 0.006031918339431286
step: 190, loss: 0.00566421914845705
step: 200, loss: 0.00032107593142427504
step: 210, loss: 0.007615870330482721
step: 220, loss: 0.004593735560774803
step: 230, loss: 0.0001492170267738402
step: 240, loss: 0.030025245621800423
step: 250, loss: 0.1128777489066124
step: 260, loss: 0.0003838404081761837
step: 270, loss: 0.00047939096111804247
step: 280, loss: 0.0034646322019398212
step: 290, loss: 0.009285996668040752
step: 300, loss: 0.030228344723582268
step: 310, loss: 0.022719528526067734
step: 320, loss: 0.007273530121892691
step: 330, loss: 0.0015149225946515799
step: 340, loss: 0.0019416025606915355
step: 350, loss: 0.031178152188658714
step: 360, loss: 0.00013638097152579576
step: 370, loss: 0.0010005461517721415
step: 380, loss: 0.015617850236594677
step: 390, loss: 0.015766749158501625
step: 400, loss: 0.0007424569339491427
step: 410, loss: 0.0004989597946405411
step: 420, loss: 0.17290237545967102
step: 430, loss: 0.0009111997205764055
step: 440, loss: 0.023218879476189613
step: 450, loss: 0.004914086312055588
step: 460, loss: 0.016708843410015106
step: 470, loss: 0.008670246228575706
step: 480, loss: 0.00863536074757576
step: 490, loss: 0.00030734832398593426
step: 500, loss: 0.0014818711206316948
step: 510, loss: 0.00040630216244608164
step: 520, loss: 0.032054826617240906

```
step: 530, loss: 0.0018861391581594944
/usr/local/lib/python3.7/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=0.899909008189263, f1=0.8978469995419147,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021433409303426743
step: 10, loss: 0.044314365833997726
step: 20, loss: 0.03997940942645073
step: 30, loss: 0.046970341354608536
step: 40, loss: 0.0005920643452554941
step: 50, loss: 0.0021575966384261847
step: 60, loss: 0.05876218155026436
step: 70, loss: 0.013223863206803799
step: 80, loss: 0.007481978740543127
step: 90, loss: 0.00047505894326604903
step: 100, loss: 0.00966248381882906
step: 110, loss: 0.004957159049808979
step: 120, loss: 0.025018561631441116
step: 130, loss: 0.009453531354665756
step: 140, loss: 0.0024902690201997757
step: 150, loss: 0.0006731484900228679
step: 160, loss: 0.0062871817499399185
step: 170, loss: 0.005796549841761589
step: 180, loss: 0.0077745346352458
step: 190, loss: 0.00048140366561710835
step: 200, loss: 0.01557796262204647
step: 210, loss: 0.0014819565694779158
step: 220, loss: 0.00024005294835660607
step: 230, loss: 0.016420090571045876
step: 240, loss: 0.004383745137602091
step: 250, loss: 0.003953835461288691
step: 260, loss: 0.0052171689458191395
step: 270, loss: 0.0001653653453104198
step: 280, loss: 0.065676644444446564
step: 290, loss: 0.046105653047561646
step: 300, loss: 0.017447300255298615
step: 310, loss: 0.0028235113713890314
step: 320, loss: 0.05373866483569145
step: 330, loss: 0.005666434299200773
step: 340, loss: 0.03998174890875816
step: 350, loss: 0.0011967809405177832
step: 360, loss: 0.004121432546526194
step: 370, loss: 0.008456064388155937
step: 380, loss: 0.017486009746789932
```

```

step: 390, loss: 0.00046587298857048154
step: 400, loss: 0.0005880692042410374
step: 410, loss: 0.0011846363777294755
step: 420, loss: 0.010603216476738453
step: 430, loss: 0.00017092627240344882
step: 440, loss: 0.000841533939819783
step: 450, loss: 0.008969760499894619
step: 460, loss: 0.0014922949485480785
step: 470, loss: 0.0582762137055397
step: 480, loss: 0.004470540676265955
step: 490, loss: 0.0026400163769721985
step: 500, loss: 0.0028206785209476948
step: 510, loss: 0.021875031292438507
step: 520, loss: 0.0037424336187541485
step: 530, loss: 0.003477084916085005
/usr/local/lib/python3.7/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=0.9052823315118397, f1=0.9097572148419605,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003604330588132143
step: 10, loss: 0.001914999564178288
step: 20, loss: 0.0018762118415907025
step: 30, loss: 0.015939688310027122
step: 40, loss: 0.0017266436479985714
step: 50, loss: 0.005586656276136637
step: 60, loss: 0.005403082352131605
step: 70, loss: 0.010665263049304485
step: 80, loss: 0.00010657250822987407
step: 90, loss: 0.020629970356822014
step: 100, loss: 0.0006050475058145821
step: 110, loss: 0.0010061990469694138
step: 120, loss: 0.006647407542914152
step: 130, loss: 0.0004817879816982895
step: 140, loss: 0.0001626188459340483
step: 150, loss: 0.0003202022926416248
step: 160, loss: 0.00022031538537703454
step: 170, loss: 0.018774211406707764
step: 180, loss: 0.00041325672646053135
step: 190, loss: 8.28747870400548e-05
step: 200, loss: 0.0015831559430807829
step: 210, loss: 0.00021092704264447093
step: 220, loss: 0.1458042412996292
step: 230, loss: 0.00084566546138376
step: 240, loss: 0.0004830112447962165

```

```

step: 250, loss: 0.009011105634272099
step: 260, loss: 0.2489132583141327
step: 270, loss: 0.00014338555047288537
step: 280, loss: 0.001528468681499362
step: 290, loss: 0.016336362808942795
step: 300, loss: 0.00017025253328029066
step: 310, loss: 0.04000135883688927
step: 320, loss: 0.0007228123140521348
step: 330, loss: 0.0038864081725478172
step: 340, loss: 0.0005912721389904618
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 350, loss: 0.13799941539764404
step: 360, loss: 0.010053356178104877
step: 370, loss: 0.0008511913474649191
step: 380, loss: 0.0007857619784772396
step: 390, loss: 0.1268918514251709
step: 400, loss: 0.03655775636434555
step: 410, loss: 0.006524866912513971
step: 420, loss: 0.0012645791284739971
step: 430, loss: 0.0040789819322526455
step: 440, loss: 0.00022647839796263725
step: 450, loss: 7.139748049667105e-05
step: 460, loss: 0.0004927537520416081
step: 470, loss: 0.0003890944644808769
step: 480, loss: 6.254246545722708e-05
step: 490, loss: 0.0001474253658670932
step: 500, loss: 0.00021129984816070646
step: 510, loss: 0.00016216399671975523
step: 520, loss: 0.00013022460916545242
step: 530, loss: 0.0010938529158011079
/usr/local/lib/python3.7/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=0.9064748201438849, f1=0.9150561797752809,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00869845226407051
step: 10, loss: 0.004987568594515324
step: 20, loss: 0.00228679901920259
step: 30, loss: 0.0003358441754244268
step: 40, loss: 0.11915326118469238
step: 50, loss: 0.00186845101416111
step: 60, loss: 0.0004092276794835925
step: 70, loss: 0.14994050562381744
step: 80, loss: 0.0240783654153347
step: 90, loss: 0.0027769014704972506

```



```
step: 100, loss: 0.00017140443378593773
step: 110, loss: 0.0006863019661977887
step: 120, loss: 0.0022772331722080708
step: 130, loss: 0.00506069278344512
step: 140, loss: 0.015560140833258629
step: 150, loss: 0.000815150560811162
step: 160, loss: 0.0012526947539299726
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.0954364687204361
step: 180, loss: 0.0003324779390823096
step: 190, loss: 0.003040899755433202
step: 200, loss: 0.003333168337121606
step: 210, loss: 0.00016209724708460271
step: 220, loss: 0.0007687131292186677
step: 230, loss: 0.0002798576606437564
step: 240, loss: 0.000656984921079129
step: 250, loss: 0.008192620240151882
step: 260, loss: 0.001422201399691403
step: 270, loss: 0.001885328209027648
step: 280, loss: 0.0024101233575493097
step: 290, loss: 0.00012687899288721383
step: 300, loss: 0.027650829404592514
step: 310, loss: 0.002548397285863757
step: 320, loss: 0.000487245328258723
step: 330, loss: 0.0015916948905214667
step: 340, loss: 0.042734451591968536
step: 350, loss: 0.0011314679868519306
step: 360, loss: 0.00686222217023373
step: 370, loss: 0.00048199607408605516
step: 380, loss: 0.0004271096258889884
step: 390, loss: 0.003075001295655966
step: 400, loss: 0.0046641710214316845
step: 410, loss: 0.0030054624658077955
step: 420, loss: 0.00011180651199538261
step: 430, loss: 0.015266257338225842
step: 440, loss: 0.00015113542031031102
step: 450, loss: 0.0004881584318354726
step: 460, loss: 0.000590158102568239
step: 470, loss: 0.00042299399501644075
step: 480, loss: 0.00018008369079325348
step: 490, loss: 0.0008770894492045045
step: 500, loss: 0.00032880809158086777
step: 510, loss: 0.0008286168449558318
step: 520, loss: 0.00021833827486261725
step: 530, loss: 0.00023327121743932366
/usr/local/lib/python3.7/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=0.9061032863849765, f1=0.9050751879699249,  
best_f1=0.9130841121495327  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0024464272428303957  
step: 10, loss: 0.0004323330067563802  
step: 20, loss: 0.0059847538359463215  
step: 30, loss: 4.7760091547388583e-05  
step: 40, loss: 0.0001427611568942666  
step: 50, loss: 0.0002455258509144187  
step: 60, loss: 0.000859182036947459  
step: 70, loss: 0.006103787571191788  
step: 80, loss: 0.0003025081241503358  
step: 90, loss: 0.0650409683585167  
step: 100, loss: 0.025034859776496887  
step: 110, loss: 0.0009026595507748425  
step: 120, loss: 0.017293429002165794  
step: 130, loss: 0.00044375413563102484  
step: 140, loss: 0.001184384454973042  
step: 150, loss: 0.00042376836063340306  
step: 160, loss: 0.0005465513095259666  
step: 170, loss: 0.0007677857647649944  
step: 180, loss: 0.0014544138684868813  
step: 190, loss: 0.0052075013518333435  
step: 200, loss: 0.0008780978387221694  
step: 210, loss: 0.001631726510822773  
step: 220, loss: 0.0005833813338540494  
step: 230, loss: 0.00047553147305734456  
step: 240, loss: 0.0019897585734725  
step: 250, loss: 0.014232922345399857  
step: 260, loss: 0.0033125211484730244  
step: 270, loss: 0.0027932035736739635  
step: 280, loss: 5.909648825763725e-05  
step: 290, loss: 0.00010826578363776207  
step: 300, loss: 0.006627128459513187  
step: 310, loss: 0.0002690217806957662  
step: 320, loss: 0.0001775723503669724  
step: 330, loss: 0.0003867439227178693  
step: 340, loss: 0.0010177802760154009  
step: 350, loss: 0.0004530487349256873  
step: 360, loss: 0.020820412784814835  
step: 370, loss: 0.002252261620014906  
step: 380, loss: 0.00023141896235756576  
step: 390, loss: 0.0008199449512176216  
step: 400, loss: 0.0022160231601446867  
step: 410, loss: 0.0002151815569959581  
step: 420, loss: 0.014644504524767399
```

```
step: 430, loss: 0.004083740990608931
step: 440, loss: 6.856050458736718e-05
step: 450, loss: 0.000964797509368509
step: 460, loss: 0.0011501394910737872
step: 470, loss: 9.978721209336072e-05
step: 480, loss: 0.0008509198087267578
step: 490, loss: 0.012986334040760994
step: 500, loss: 0.0061937193386256695
step: 510, loss: 0.0006319183739833534
step: 520, loss: 0.0015883267624303699
step: 530, loss: 0.0006377323297783732
/usr/local/lib/python3.7/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=0.9034038638454462, f1=0.9092592592592593,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0056946761906147
step: 10, loss: 0.1926327645778656
step: 20, loss: 0.0005190170486457646
step: 30, loss: 0.05040840804576874
step: 40, loss: 0.004256885498762131
step: 50, loss: 0.0007708861376158893
step: 60, loss: 0.0010188784217461944
step: 70, loss: 0.00043590442510321736
step: 80, loss: 0.00016596316709183156
step: 90, loss: 0.00046695026685483754
step: 100, loss: 0.0013443591305986047
step: 110, loss: 0.0007475549355149269
step: 120, loss: 0.006090659648180008
step: 130, loss: 0.00039096103864721954
step: 140, loss: 0.00012441213766578585
step: 150, loss: 0.00017597494297660887
step: 160, loss: 0.0014919493114575744
step: 170, loss: 0.008039653301239014
step: 180, loss: 0.00022669554164167494
step: 190, loss: 0.0005315967719070613
step: 200, loss: 0.00023545691510662436
step: 210, loss: 0.0005698928143829107
step: 220, loss: 0.0010049936827272177
step: 230, loss: 0.0010373928816989064
step: 240, loss: 0.0001635501248529181
step: 250, loss: 0.00016622037219349295
step: 260, loss: 0.00015841245476622134
step: 270, loss: 0.00022211577743291855
step: 280, loss: 0.001188037684187293
```

```

step: 290, loss: 0.00010472638678038493
step: 300, loss: 0.005777280777692795
step: 310, loss: 0.00171846360899508
step: 320, loss: 0.0005969768390059471
step: 330, loss: 0.005447596777230501
step: 340, loss: 0.006848529446870089
step: 350, loss: 0.00014144295710138977
step: 360, loss: 0.002105315448716283
step: 370, loss: 0.00012226500257384032
step: 380, loss: 3.697256033774465e-05
step: 390, loss: 0.0017861895030364394
step: 400, loss: 4.633342905435711e-05
step: 410, loss: 0.04032447189092636
step: 420, loss: 0.004256522748619318
step: 430, loss: 0.0001128736839746125
step: 440, loss: 0.0001705775794107467
step: 450, loss: 0.00016581967065576464
step: 460, loss: 0.008080530911684036
step: 470, loss: 0.000911232375074178
step: 480, loss: 0.0007747671334072948
step: 490, loss: 0.001405062386766076
step: 500, loss: 0.002222128212451935
step: 510, loss: 0.0001515358017059043
step: 520, loss: 8.976735261967406e-05
step: 530, loss: 0.0002870836469810456
/usr/local/lib/python3.7/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=0.9059040590405903, f1=0.9082526509912402,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003326302976347506
step: 10, loss: 0.0003984309150837362
step: 20, loss: 5.675682405126281e-05
step: 30, loss: 0.0031274149660021067
step: 40, loss: 0.00013595020573120564
step: 50, loss: 0.0003376683162059635
step: 60, loss: 0.005920563358813524
step: 70, loss: 0.0005421750247478485
step: 80, loss: 0.0004974052426405251
step: 90, loss: 0.10170099884271622
step: 100, loss: 0.00013870645489078015
step: 110, loss: 0.0003766522277146578
step: 120, loss: 0.0002715890295803547
step: 130, loss: 0.0006601143977604806
step: 140, loss: 0.00010653434583218768

```

```

step: 150, loss: 0.00033419724786654115
step: 160, loss: 0.00012927762873005122
step: 170, loss: 0.0028933631256222725
step: 180, loss: 8.134653035085648e-05
step: 190, loss: 0.0001310625666519627
step: 200, loss: 0.0002578525454737246
step: 210, loss: 0.00032264486071653664
step: 220, loss: 6.173945439513773e-05
step: 230, loss: 0.00032421379000879824
step: 240, loss: 0.0002572644443716854
step: 250, loss: 0.000113618079922162
step: 260, loss: 0.0005411674501374364
step: 270, loss: 0.0004774273547809571
step: 280, loss: 9.669893188402057e-05
step: 290, loss: 0.0014232207322493196
step: 300, loss: 0.19243542850017548
step: 310, loss: 0.0003121024346910417
step: 320, loss: 0.000607138907071203
step: 330, loss: 0.005885298363864422
step: 340, loss: 0.00042087852489203215
step: 350, loss: 0.0008793122251518071
step: 360, loss: 0.00028128287522122264
step: 370, loss: 0.005866260267794132
step: 380, loss: 0.00028960013878531754
step: 390, loss: 8.016101492103189e-05
step: 400, loss: 0.00019624349079094827
step: 410, loss: 0.001183749525807798
step: 420, loss: 0.0005356455803848803
step: 430, loss: 0.0006310345488600433
step: 440, loss: 0.0013662258861586452
step: 450, loss: 0.0008169006323441863
step: 460, loss: 0.00022264114522840828
step: 470, loss: 0.1828119158744812
step: 480, loss: 0.00013683876022696495
step: 490, loss: 0.0015335187781602144
step: 500, loss: 0.001958211651071906
step: 510, loss: 0.0013372163521125913
step: 520, loss: 0.0016019026516005397
step: 530, loss: 0.00020121487614233047
/usr/local/lib/python3.7/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=0.9065550906555091, f1=0.9083294555607259,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004283631860744208

```

step: 10, loss: 0.0002134100504918024
step: 20, loss: 0.00024572850088588893
step: 30, loss: 0.00012680626241490245
step: 40, loss: 0.0001999349769903347
step: 50, loss: 0.00015962582256179303
step: 60, loss: 8.045576396398246e-05
step: 70, loss: 0.00010444646613905206
step: 80, loss: 0.0006731055909767747
step: 90, loss: 0.00017312158888671547
step: 100, loss: 0.0001581391115905717
step: 110, loss: 0.000118847397970967
step: 120, loss: 0.0006293116603046656
step: 130, loss: 0.0012266197009012103
step: 140, loss: 0.0008290260448120534
step: 150, loss: 0.004748891107738018
step: 160, loss: 0.001465042121708393
step: 170, loss: 6.567943637492135e-05
step: 180, loss: 9.640550706535578e-05
step: 190, loss: 0.0007500422070734203
step: 200, loss: 0.00013613289047498256
step: 210, loss: 9.611996210878715e-05
step: 220, loss: 5.945135490037501e-05
step: 230, loss: 5.8714977058116347e-05
step: 240, loss: 3.947946970583871e-05
step: 250, loss: 0.003247275482863188
step: 260, loss: 0.00032636660034768283
step: 270, loss: 5.9017984312959015e-05
step: 280, loss: 0.00011912341869901866
step: 290, loss: 0.0001965737174032256
step: 300, loss: 0.001195093966089189
step: 310, loss: 0.00010559243673924357
step: 320, loss: 0.001568678766489029
step: 330, loss: 0.0003529712266754359
step: 340, loss: 4.968438224750571e-05
step: 350, loss: 0.01946224272251129
step: 360, loss: 4.6354802179848775e-05
step: 370, loss: 0.00013760378351435065
step: 380, loss: 0.0005368429701775312
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.0002829136501532048
step: 400, loss: 0.00010437600576551631
step: 410, loss: 0.005230704788118601
step: 420, loss: 0.001332922838628292
step: 430, loss: 0.031401127576828
step: 440, loss: 6.5598200308159e-05
step: 450, loss: 3.968052624259144e-05
step: 460, loss: 3.568704778444953e-05
step: 470, loss: 0.002443561563268304

```

step: 480, loss: 0.00017234797996934503
step: 490, loss: 0.00023253026301972568
step: 500, loss: 9.227173723047599e-05
step: 510, loss: 4.059646016685292e-05
step: 520, loss: 8.065462316153571e-05
step: 530, loss: 0.00014319655019789934
/usr/local/lib/python3.7/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=0.9107142857142857, f1=0.9032863849765258,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003664109099190682
step: 10, loss: 0.006950105540454388
step: 20, loss: 9.753408085089177e-05
step: 30, loss: 0.0002423321275273338
step: 40, loss: 0.002518374938517809
step: 50, loss: 0.00024701215443201363
step: 60, loss: 4.0614781028125435e-05
step: 70, loss: 0.00017477023357059807
step: 80, loss: 9.806977323023602e-05
step: 90, loss: 5.036469519836828e-05
step: 100, loss: 0.000665623287204653
step: 110, loss: 4.8802601668285206e-05
step: 120, loss: 0.00016325624892488122
step: 130, loss: 0.0002067700697807595
step: 140, loss: 0.008886382915079594
step: 150, loss: 0.020256027579307556
step: 160, loss: 0.00022331222135107964
step: 170, loss: 0.002625645836815238
step: 180, loss: 0.0004577519139274955
step: 190, loss: 0.00026809590053744614
step: 200, loss: 0.0006883060559630394
step: 210, loss: 0.0023194137029349804
step: 220, loss: 0.0005779642378911376
step: 230, loss: 8.786646503722295e-05
step: 240, loss: 0.000317582453135401
step: 250, loss: 6.57134223729372e-05
step: 260, loss: 7.7814758697059e-05
step: 270, loss: 0.0011527135502547026
step: 280, loss: 0.0007214222569018602
step: 290, loss: 0.0006784942233934999
step: 300, loss: 0.00022757274564355612
step: 310, loss: 5.8350142353447154e-05
step: 320, loss: 0.00027670804411172867
step: 330, loss: 9.371310443384573e-05

```

```

step: 340, loss: 0.0009564241045154631
step: 350, loss: 0.0003650347061920911
step: 360, loss: 0.00014958259998820722
step: 370, loss: 0.00013191184552852064
step: 380, loss: 0.17582985758781433
step: 390, loss: 0.00129637960344553
step: 400, loss: 0.000182959163794294
step: 410, loss: 0.0018988418160006404
step: 420, loss: 0.0006359688704833388
step: 430, loss: 0.0006797884125262499
step: 440, loss: 0.0007480567437596619
step: 450, loss: 0.0002064177388092503
step: 460, loss: 0.0008161197765730321
step: 470, loss: 0.0035526209976524115
step: 480, loss: 0.00011084898869739845
step: 490, loss: 0.0006197378970682621
step: 500, loss: 0.00048222290934063494
step: 510, loss: 0.00013608267181552947
step: 520, loss: 8.30835269880481e-05
step: 530, loss: 0.00042301134089939296
/usr/local/lib/python3.7/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=0.9069097888675625, f1=0.8946360153256706,
best_f1=0.9130841121495327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027309171855449677
step: 10, loss: 0.000564929039683193
step: 20, loss: 0.0002934546791948378
step: 30, loss: 0.00036173389526084065
step: 40, loss: 0.00013832339027430862
step: 50, loss: 0.0005580642609857023
step: 60, loss: 0.0006916156271472573
step: 70, loss: 5.5491396778961644e-05
step: 80, loss: 0.0008457742515020072
step: 90, loss: 0.00011139762500533834
step: 100, loss: 5.4690131946699694e-05
step: 110, loss: 0.0012740498641505837
step: 120, loss: 0.00037782787694595754
step: 130, loss: 8.362218795809895e-05
step: 140, loss: 0.00011611756781348959
step: 150, loss: 0.01432315818965435
step: 160, loss: 0.0103230532258749
step: 170, loss: 0.0002118548727594316
step: 180, loss: 0.00036855266080237925
step: 190, loss: 0.0022796366829425097

```



```

step: 200, loss: 0.00010014015424530953
step: 210, loss: 0.00012303346011321992
step: 220, loss: 7.19880627002567e-05
step: 230, loss: 0.000986485625617206
step: 240, loss: 0.000967320054769516
step: 250, loss: 0.0002761392679531127
step: 260, loss: 0.0009135493892244995
step: 270, loss: 0.00018852944776881486
step: 280, loss: 0.00033620846807025373
step: 290, loss: 0.00038586760638281703
step: 300, loss: 0.0020323197823017836
step: 310, loss: 0.004490051884204149
step: 320, loss: 0.00012397811224218458
step: 330, loss: 0.00015624091611243784
step: 340, loss: 0.0006898603169247508
step: 350, loss: 0.0005663100746460259
step: 360, loss: 0.001898624119348824
step: 370, loss: 0.0002118181873811409
step: 380, loss: 5.950099148321897e-05
step: 390, loss: 0.0012609277619048953
step: 400, loss: 5.0635284424060956e-05
step: 410, loss: 0.00015865737805143
step: 420, loss: 0.00016982028319034725
step: 430, loss: 7.906230894150212e-05
step: 440, loss: 0.0001562916731927544
step: 450, loss: 0.000957137846853584
step: 460, loss: 0.00022938389156479388
step: 470, loss: 0.0001244789600605145
step: 480, loss: 0.00023904209956526756
step: 490, loss: 0.0001390296092722565
step: 500, loss: 7.505285611841828e-05
step: 510, loss: 0.00033278571208938956
step: 520, loss: 0.00012848323967773467
step: 530, loss: 0.00040454111876897514
/usr/local/lib/python3.7/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=0.9069212410501194, f1=0.8968481375358165,
best_f1=0.9130841121495327
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 [00:16, 354.43it/s]
load_f1 = 0.9074598677998111
real_f1 = 0.9054117647058824
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 [00:12, 360.05it/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 distilbert \
        --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 distilbert \
--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: 0.8446950316429138

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

step: 10, loss: 0.0713365226984024

step: 20, loss: 0.3859333097934723

step: 30, loss: 0.36956316232681274

step: 40, loss: 0.5167160034179688

step: 50, loss: 0.3085872530937195

step: 60, loss: 0.37301695346832275

step: 70, loss: 0.25763699412345886

step: 80, loss: 0.26277756690979004

step: 90, loss: 0.4513276517391205

step: 100, loss: 0.19628550112247467

step: 110, loss: 0.3103722631931305

step: 120, loss: 0.25245946645736694

step: 130, loss: 0.2276798039674759

step: 140, loss: 0.28783100843429565

step: 150, loss: 0.28802213072776794

step: 160, loss: 0.2422104775905609

step: 170, loss: 0.18222860991954803

step: 180, loss: 0.18905773758888245

step: 190, loss: 0.24738715589046478

step: 200, loss: 0.17818138003349304

step: 210, loss: 0.469903826713562

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

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

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

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=0.488245931283906, f1=0.5335689045936396,

best_f1=0.5335689045936396

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

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

step: 0, loss: 0.07110513001680374

step: 10, loss: 0.08665168285369873

step: 20, loss: 0.30568286776542664

step: 30, loss: 0.2676050662994385

step: 40, loss: 0.0779406800866127

```

step: 50, loss: 0.17558708786964417
step: 60, loss: 0.09312599152326584
step: 70, loss: 0.28202247619628906
step: 80, loss: 0.25012391805648804
step: 90, loss: 0.17568805813789368
step: 100, loss: 0.13590963184833527
step: 110, loss: 0.06965599954128265
step: 120, loss: 0.2014811784029007
step: 130, loss: 0.2527758479118347
step: 140, loss: 0.278944194316864
step: 150, loss: 0.2975308299064636
step: 160, loss: 0.13001784682273865
step: 170, loss: 0.21828213334083557
step: 180, loss: 0.2226148396730423
step: 190, loss: 0.1811293363571167
step: 200, loss: 0.1272081583738327
step: 210, loss: 0.46597936749458313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5612648221343873, f1=0.5488721804511278,
best_f1=0.5488721804511278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24124282598495483
step: 10, loss: 0.2666139006614685
step: 20, loss: 0.2674378752708435
step: 30, loss: 0.08415556699037552
step: 40, loss: 0.1592424064874649
step: 50, loss: 0.20814776420593262
step: 60, loss: 0.17615561187267303
step: 70, loss: 0.08662450313568115
step: 80, loss: 0.07411877810955048
step: 90, loss: 0.16995368897914886
step: 100, loss: 0.10968464612960815
step: 110, loss: 0.17016823589801788
step: 120, loss: 0.13888025283813477
step: 130, loss: 0.18011754751205444
step: 140, loss: 0.22173379361629486
step: 150, loss: 0.3609728515148163
step: 160, loss: 0.12848348915576935
step: 170, loss: 0.2934972941875458
step: 180, loss: 0.1791096031665802

```

```

step: 190, loss: 0.24290448427200317
step: 200, loss: 0.10764417797327042
step: 210, loss: 0.2418275773525238
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5803571428571429, f1=0.5279642058165548,
best_f1=0.5279642058165548
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18389059603214264
step: 10, loss: 0.13681761920452118
step: 20, loss: 0.20915545523166656
step: 30, loss: 0.0633172020316124
step: 40, loss: 0.15073423087596893
step: 50, loss: 0.06052277609705925
step: 60, loss: 0.04489060863852501
step: 70, loss: 0.23840779066085815
step: 80, loss: 0.20623226463794708
step: 90, loss: 0.052144668996334076
step: 100, loss: 0.12044200301170349
step: 110, loss: 0.1825796365737915
step: 120, loss: 0.09106070548295975
step: 130, loss: 0.11607787758111954
step: 140, loss: 0.30965498089790344
step: 150, loss: 0.2974552810192108
step: 160, loss: 0.06642799079418182
step: 170, loss: 0.09547340124845505
step: 180, loss: 0.09853511303663254
step: 190, loss: 0.19351357221603394
step: 200, loss: 0.17982414364814758
step: 210, loss: 0.03719126805663109
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5809312638580932, f1=0.5256124721603563,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.070970818400383
step: 10, loss: 0.07200340181589127
step: 20, loss: 0.07046245038509369
step: 30, loss: 0.229170560836792
step: 40, loss: 0.21702918410301208
step: 50, loss: 0.25281691551208496
step: 60, loss: 0.08594948798418045
step: 70, loss: 0.20232753455638885
step: 80, loss: 0.039720334112644196
step: 90, loss: 0.08906558156013489
step: 100, loss: 0.08325318247079849
step: 110, loss: 0.05505802482366562
step: 120, loss: 0.11528217047452927
step: 130, loss: 0.1100437119603157
step: 140, loss: 0.08269945532083511
step: 150, loss: 0.1239078938961029
step: 160, loss: 0.07711897790431976
step: 170, loss: 0.18518100678920746
step: 180, loss: 0.3102339804172516
step: 190, loss: 0.1292383074760437
step: 200, loss: 0.11064902693033218
step: 210, loss: 0.04872506856918335
/usr/local/lib/python3.7/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=0.5349650349650349, f1=0.5274336283185841,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06394539028406143
step: 10, loss: 0.03208897262811661
step: 20, loss: 0.10181967169046402
step: 30, loss: 0.3579655885696411
step: 40, loss: 0.12445393204689026
step: 50, loss: 0.029315723106265068
step: 60, loss: 0.07916216552257538
step: 70, loss: 0.007237857207655907
step: 80, loss: 0.0945575088262558
step: 90, loss: 0.10494376718997955
step: 100, loss: 0.0727679505944252
step: 110, loss: 0.027129730209708214
step: 120, loss: 0.05079704523086548
step: 130, loss: 0.075562484562397
step: 140, loss: 0.11892203241586685
step: 150, loss: 0.023072602227330208
step: 160, loss: 0.15381300449371338

```

```

step: 170, loss: 0.04497702419757843
step: 180, loss: 0.01560913771390915
step: 190, loss: 0.07717655599117279
step: 200, loss: 0.028885971754789352
step: 210, loss: 0.004607593175023794
/usr/local/lib/python3.7/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=0.5450450450450449, f1=0.5136363636363637,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0536489263176918
step: 10, loss: 0.07216981053352356
step: 20, loss: 0.042556099593639374
step: 30, loss: 0.028880923986434937
step: 40, loss: 0.1183147057890892
step: 50, loss: 0.10154860466718674
step: 60, loss: 0.20955413579940796
step: 70, loss: 0.014806010760366917
step: 80, loss: 0.016818011179566383
step: 90, loss: 0.003005731152370572
step: 100, loss: 0.005236377939581871
step: 110, loss: 0.2309715300798416
step: 120, loss: 0.18317978084087372
step: 130, loss: 0.009789896197617054
step: 140, loss: 0.002952825976535678
step: 150, loss: 0.018295319750905037
step: 160, loss: 0.12251371890306473
step: 170, loss: 0.14162062108516693
step: 180, loss: 0.017631495371460915
step: 190, loss: 0.046401552855968475
step: 200, loss: 0.0684046670794487
step: 210, loss: 0.048773981630802155
/usr/local/lib/python3.7/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=0.5510204081632653, f1=0.5407554671968191,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04657173529267311
step: 10, loss: 0.004829325247555971
step: 20, loss: 0.041874539107084274
step: 30, loss: 0.05553501099348068
step: 40, loss: 0.01623542234301567

```



```

step: 50, loss: 0.03902590274810791
step: 60, loss: 0.23994776606559753
step: 70, loss: 0.003920888062566519
step: 80, loss: 0.004323990549892187
step: 90, loss: 0.23141291737556458
step: 100, loss: 0.050649404525756836
step: 110, loss: 0.007150455377995968
step: 120, loss: 0.041039370000362396
step: 130, loss: 0.04049617424607277
step: 140, loss: 0.029740888625383377
step: 150, loss: 0.07576625794172287
step: 160, loss: 0.1258254498243332
step: 170, loss: 0.07751492410898209
step: 180, loss: 0.014017673209309578
step: 190, loss: 0.005031915847212076
step: 200, loss: 0.07954450696706772
step: 210, loss: 0.20593896508216858
/usr/local/lib/python3.7/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=0.543046357615894, f1=0.5387931034482758,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018129024654626846
step: 10, loss: 0.029626015573740005
step: 20, loss: 0.1749391406774521
step: 30, loss: 0.37604543566703796
step: 40, loss: 0.05389988794922829
step: 50, loss: 0.007911447435617447
step: 60, loss: 0.22313182055950165
step: 70, loss: 0.11773066222667694
step: 80, loss: 0.00618447782471776
step: 90, loss: 0.06281294673681259
step: 100, loss: 0.007500025909394026
step: 110, loss: 0.007228297181427479
step: 120, loss: 0.2248689979314804
step: 130, loss: 0.012010888196527958
step: 140, loss: 0.05516871064901352
step: 150, loss: 0.044964421540498734
step: 160, loss: 0.003520873375236988
step: 170, loss: 0.10704827308654785
step: 180, loss: 0.012477420270442963
step: 190, loss: 0.10783708840608597
step: 200, loss: 0.021222319453954697
step: 210, loss: 0.07858218252658844
/usr/local/lib/python3.7/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=0.540650406504065, f1=0.524, best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02409411035478115
step: 10, loss: 0.0030189601238816977
step: 20, loss: 0.0656290054321289
step: 30, loss: 0.29373759031295776
step: 40, loss: 0.12435662001371384
step: 50, loss: 0.0072032916359603405
step: 60, loss: 0.0056697530671954155
step: 70, loss: 0.06820205599069595
step: 80, loss: 0.008892412297427654
step: 90, loss: 0.023743802681565285
step: 100, loss: 0.014474324882030487
step: 110, loss: 0.008298303000628948
step: 120, loss: 0.02709255926311016
step: 130, loss: 0.029832465574145317
step: 140, loss: 0.001265055499970913
step: 150, loss: 0.16421978175640106
step: 160, loss: 0.0717906802892685
step: 170, loss: 0.07845667004585266
step: 180, loss: 0.04335218295454979
step: 190, loss: 0.0399734228849411
step: 200, loss: 0.014207876287400723
step: 210, loss: 0.030546752735972404
/usr/local/lib/python3.7/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=0.5384615384615384, f1=0.5392354124748492,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012286253273487091
step: 10, loss: 0.018269959837198257
step: 20, loss: 0.10651785880327225
step: 30, loss: 0.04514986649155617
step: 40, loss: 0.05175522714853287
step: 50, loss: 0.07738308608531952
step: 60, loss: 0.008892043493688107
step: 70, loss: 0.014524701051414013
step: 80, loss: 0.011093036271631718
step: 90, loss: 0.11404909193515778
step: 100, loss: 0.05466224625706673
step: 110, loss: 0.09200811386108398

```

```

step: 120, loss: 0.07885375618934631
step: 130, loss: 0.06574759632349014
step: 140, loss: 0.04761599749326706
step: 150, loss: 0.031978070735931396
step: 160, loss: 0.0020731117110699415
step: 170, loss: 0.26683685183525085
step: 180, loss: 0.06987110525369644
step: 190, loss: 0.04231458902359009
step: 200, loss: 0.0028546496760100126
step: 210, loss: 0.0036871435586363077
/usr/local/lib/python3.7/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=0.5378787878787878, f1=0.525328330206379,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021442861761897802
step: 10, loss: 0.008360027335584164
step: 20, loss: 0.009411882609128952
step: 30, loss: 0.006231597624719143
step: 40, loss: 0.28847405314445496
step: 50, loss: 0.0677495151758194
step: 60, loss: 0.010967721231281757
step: 70, loss: 0.00268289796076715
step: 80, loss: 0.05601923540234566
step: 90, loss: 0.08003939688205719
step: 100, loss: 0.04168625921010971
step: 110, loss: 0.004202916286885738
step: 120, loss: 0.0018273640889674425
step: 130, loss: 0.00900351069867611
step: 140, loss: 0.0836474746465683
step: 150, loss: 0.0037131512071937323
step: 160, loss: 0.0228776466101408
step: 170, loss: 0.016279637813568115
step: 180, loss: 0.01021566055715084
step: 190, loss: 0.22366096079349518
step: 200, loss: 0.10432624071836472
step: 210, loss: 0.029822228476405144
/usr/local/lib/python3.7/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=0.5387596899224806, f1=0.5160075329566856,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0013280619168654084
step: 10, loss: 0.019461145624518394
step: 20, loss: 0.005968020297586918
step: 30, loss: 0.028819916769862175
step: 40, loss: 0.0028818233404308558
step: 50, loss: 0.0011301393387839198
step: 60, loss: 0.01218991819769144
step: 70, loss: 0.022511858493089676
step: 80, loss: 0.12713560461997986
step: 90, loss: 0.0027759491931647062
step: 100, loss: 0.04181969538331032
step: 110, loss: 0.013692074455320835
step: 120, loss: 0.0013794696424156427
step: 130, loss: 0.020266858860850334
step: 140, loss: 0.0648154765367508
step: 150, loss: 0.002479179762303829
step: 160, loss: 0.0033740338403731585
step: 170, loss: 0.0013703972799703479
step: 180, loss: 0.026878956705331802
step: 190, loss: 0.0025398393627256155
step: 200, loss: 0.012645374052226543
step: 210, loss: 0.010566000826656818
/usr/local/lib/python3.7/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=0.5405405405405406, f1=0.5123339658444023,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013207473792135715
step: 10, loss: 0.0015511811943724751
step: 20, loss: 0.0013385581551119685
step: 30, loss: 0.0008390287403017282
step: 40, loss: 0.003104698145762086
step: 50, loss: 0.0042717792093753815
step: 60, loss: 0.0009737984510138631
step: 70, loss: 0.05711614340543747
step: 80, loss: 0.06360892951488495
step: 90, loss: 0.28079041838645935
step: 100, loss: 0.005449655465781689
step: 110, loss: 0.006101544480770826
step: 120, loss: 0.008706677705049515
step: 130, loss: 0.00269026099704206
step: 140, loss: 0.0011429006699472666
step: 150, loss: 0.03043806180357933
step: 160, loss: 0.002626601606607437
step: 170, loss: 0.014291031286120415

```

```

step: 180, loss: 0.005329383537173271
step: 190, loss: 0.008224652148783207
step: 200, loss: 0.006864238064736128
step: 210, loss: 0.008842711336910725
/usr/local/lib/python3.7/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=0.530938123752495, f1=0.503968253968254,
best_f1=0.5256124721603563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000880870851688087
step: 10, loss: 0.03365721181035042
step: 20, loss: 0.004538642708212137
step: 30, loss: 0.0009158928878605366
step: 40, loss: 0.0008149007917381823
step: 50, loss: 0.0028272520285099745
step: 60, loss: 0.04085945710539818
step: 70, loss: 0.0034728029277175665
step: 80, loss: 0.11882974207401276
step: 90, loss: 0.0428863987326622
step: 100, loss: 0.17746266722679138
step: 110, loss: 0.0009478600695729256
step: 120, loss: 0.026322554796934128
step: 130, loss: 0.0033851354382932186
step: 140, loss: 0.0015973594272509217
step: 150, loss: 0.014553096145391464
step: 160, loss: 0.0019319580169394612
step: 170, loss: 0.03381756320595741
step: 180, loss: 0.06351771950721741
step: 190, loss: 0.018584270030260086
step: 200, loss: 0.0265522338449955
step: 210, loss: 0.0474809892475605
/usr/local/lib/python3.7/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=0.5335968379446641, f1=0.5048923679060666,
best_f1=0.5256124721603563
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, 546.98it/s]
load_f1 = 0.5805084745762712
real_f1 = 0.5751633986928104
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 [00:12, 361.91it/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 distilbert \
  --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 distilbert \
  --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: 0.8441784977912903

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

step: 10, loss: 0.1626245081424713

step: 20, loss: 0.1536642462015152

step: 30, loss: 0.5088537931442261

step: 40, loss: 0.2590026259422302

step: 50, loss: 0.307249516248703

step: 60, loss: 0.3694600462913513

step: 70, loss: 0.179440438747406

step: 80, loss: 0.5190481543540955

step: 90, loss: 0.23637086153030396

step: 100, loss: 0.22298845648765564

step: 110, loss: 0.24256831407546997

step: 120, loss: 0.42062416672706604

step: 130, loss: 0.34956130385398865

step: 140, loss: 0.33657926321029663

step: 150, loss: 0.2707206904888153

step: 160, loss: 0.22043192386627197

step: 170, loss: 0.3919826149940491

step: 180, loss: 0.27818647027015686

step: 190, loss: 0.15383146703243256

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

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

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

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=0.5012531328320802, f1=0.5405405405405406,

best_f1=0.5405405405405406

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

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

step: 0, loss: 0.3595565855503082

step: 10, loss: 0.0353207141160965

step: 20, loss: 0.09700517356395721

step: 30, loss: 0.15444466471672058

step: 40, loss: 0.4783799648284912

step: 50, loss: 0.29636067152023315

step: 60, loss: 0.16423188149929047


```

step: 70, loss: 0.287867933511734
step: 80, loss: 0.2033742368221283
step: 90, loss: 0.1283348649740219
step: 100, loss: 0.23380893468856812
step: 110, loss: 0.13496306538581848
step: 120, loss: 0.18233121931552887
step: 130, loss: 0.055441077798604965
step: 140, loss: 0.22151239216327667
step: 150, loss: 0.031070562079548836
step: 160, loss: 0.10982048511505127
step: 170, loss: 0.13252393901348114
step: 180, loss: 0.22815123200416565
step: 190, loss: 0.17360028624534607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7208672086720868, f1=0.7210526315789473,
best_f1=0.7210526315789473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1438167691230774
step: 10, loss: 0.1839408576488495
step: 20, loss: 0.052355971187353134
step: 30, loss: 0.04730171337723732
step: 40, loss: 0.061347946524620056
step: 50, loss: 0.2384171038866043
step: 60, loss: 0.14829209446907043
step: 70, loss: 0.1713469922542572
step: 80, loss: 0.12631691992282867
step: 90, loss: 0.17643897235393524
step: 100, loss: 0.15516185760498047
step: 110, loss: 0.22104333341121674
step: 120, loss: 0.015834521502256393
step: 130, loss: 0.035664837807416916
step: 140, loss: 0.13911207020282745
step: 150, loss: 0.12413179129362106
step: 160, loss: 0.0883711650967598
step: 170, loss: 0.035495560616254807
step: 180, loss: 0.09111844748258591
step: 190, loss: 0.21705782413482666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7278106508875739, f1=0.7714285714285715,
best_f1=0.7714285714285715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14586669206619263
step: 10, loss: 0.09278328716754913
step: 20, loss: 0.025141187012195587
step: 30, loss: 0.04314856603741646
step: 40, loss: 0.09928231686353683
step: 50, loss: 0.0801321491599083
step: 60, loss: 0.1076088547706604
step: 70, loss: 0.04615813493728638
step: 80, loss: 0.0346180722117424
step: 90, loss: 0.04966707527637482
step: 100, loss: 0.07599318772554398
step: 110, loss: 0.09224182367324829
step: 120, loss: 0.06017833575606346
step: 130, loss: 0.14325334131717682
step: 140, loss: 0.13910488784313202
step: 150, loss: 0.011438140645623207
step: 160, loss: 0.0962962955236435
step: 170, loss: 0.09658913314342499
step: 180, loss: 0.13429461419582367
step: 190, loss: 0.09065829962491989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7425474254742548, f1=0.7616580310880829,
best_f1=0.7616580310880829
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02051231637597084
step: 10, loss: 0.0447876900434494
step: 20, loss: 0.13151690363883972
step: 30, loss: 0.1008947417140007
step: 40, loss: 0.031377583742141724
step: 50, loss: 0.01847071573138237
step: 60, loss: 0.06863076239824295

```

```

step: 70, loss: 0.010188479907810688
step: 80, loss: 0.008869373239576817
step: 90, loss: 0.0013357471907511353
step: 100, loss: 0.06278304755687714
step: 110, loss: 0.004055027849972248
step: 120, loss: 0.011372790671885014
step: 130, loss: 0.10290312767028809
step: 140, loss: 0.007120088208466768
step: 150, loss: 0.09760889410972595
step: 160, loss: 0.03872689977288246
step: 170, loss: 0.04729727655649185
step: 180, loss: 0.10314854234457016
step: 190, loss: 0.17829404771327972
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7430167597765364, f1=0.7814207650273224,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16307252645492554
step: 10, loss: 0.005934087093919516
step: 20, loss: 0.018159106373786926
step: 30, loss: 0.13168001174926758
step: 40, loss: 0.03781450167298317
step: 50, loss: 0.004193218424916267
step: 60, loss: 0.004760525655001402
step: 70, loss: 0.14408248662948608
step: 80, loss: 0.1399819701910019
step: 90, loss: 0.032159142196178436
step: 100, loss: 0.028786689043045044
step: 110, loss: 0.21722663938999176
step: 120, loss: 0.07239760458469391
step: 130, loss: 0.012261902913451195
step: 140, loss: 0.009028846397995949
step: 150, loss: 0.058393239974975586
step: 160, loss: 0.020914118736982346
step: 170, loss: 0.008024183101952076
step: 180, loss: 0.0036629708483815193
step: 190, loss: 0.04657981917262077
/usr/local/lib/python3.7/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=0.7311827956989247, f1=0.768421052631579,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017514569917693734
step: 10, loss: 0.012640646658837795
step: 20, loss: 0.009284872561693192
step: 30, loss: 0.07918087393045425
step: 40, loss: 0.016838952898979187
step: 50, loss: 0.15714405477046967
step: 60, loss: 0.0266161747276783
step: 70, loss: 0.006883684080094099
step: 80, loss: 0.012096989899873734
step: 90, loss: 0.1098148301243782
step: 100, loss: 0.01677051931619644
step: 110, loss: 0.0036048090551048517
step: 120, loss: 0.023438388481736183
step: 130, loss: 0.012059995904564857
step: 140, loss: 0.005939680151641369
step: 150, loss: 0.010713196359574795
step: 160, loss: 0.0036337501369416714
step: 170, loss: 0.030291331931948662
step: 180, loss: 0.003061749739572406
step: 190, loss: 0.005941013339906931
/usr/local/lib/python3.7/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=0.7356948228882835, f1=0.7849462365591398,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002159939846023917
step: 10, loss: 0.1117606833577156
step: 20, loss: 0.0010968113783746958
step: 30, loss: 0.055021923035383224
step: 40, loss: 0.030574830248951912
step: 50, loss: 0.003101983806118369
step: 60, loss: 0.06588895618915558
step: 70, loss: 0.007731957361102104
step: 80, loss: 0.058653298765420914
step: 90, loss: 0.0021001354325562716
step: 100, loss: 0.03270857036113739
step: 110, loss: 0.006089040543884039
step: 120, loss: 0.0028733101207762957
step: 130, loss: 0.0019491874845698476
step: 140, loss: 0.028854284435510635

```

```

step: 150, loss: 0.1783069372177124
step: 160, loss: 0.014969699084758759
step: 170, loss: 0.011806043796241283
step: 180, loss: 0.047859013080596924
step: 190, loss: 0.06865553557872772
/usr/local/lib/python3.7/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=0.7283582089552239, f1=0.7596439169139465,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02836163528263569
step: 10, loss: 0.01792292110621929
step: 20, loss: 0.03632911294698715
step: 30, loss: 0.13401368260383606
step: 40, loss: 0.029069356620311737
step: 50, loss: 0.0020198451820760965
step: 60, loss: 0.0017702269833534956
step: 70, loss: 0.0008540436392650008
step: 80, loss: 0.0024768763687461615
step: 90, loss: 0.008337646722793579
step: 100, loss: 0.0028839753940701485
step: 110, loss: 0.002487601013854146
step: 120, loss: 0.21905218064785004
step: 130, loss: 0.001622454379685223
step: 140, loss: 0.0016907122917473316
step: 150, loss: 0.010055065155029297
step: 160, loss: 0.022389981895685196
step: 170, loss: 0.1020364910364151
step: 180, loss: 0.21208684146404266
step: 190, loss: 0.0434422492980957
/usr/local/lib/python3.7/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=0.7028571428571428, f1=0.7500000000000001,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005953222280368209
step: 10, loss: 0.07554441690444946
step: 20, loss: 0.0019135131733492017
step: 30, loss: 0.0020360909402370453
step: 40, loss: 0.00036858581006526947
step: 50, loss: 0.0010077713523060083
step: 60, loss: 0.06330744922161102

```

```

step: 70, loss: 0.0036812792532145977
step: 80, loss: 0.0007043805671855807
step: 90, loss: 0.036601126194000244
step: 100, loss: 0.016231583431363106
step: 110, loss: 0.005122964736074209
step: 120, loss: 0.0199415385723114
step: 130, loss: 0.13323312997817993
step: 140, loss: 0.0012689523864537477
step: 150, loss: 0.001179794780910015
step: 160, loss: 0.0007570774178020656
step: 170, loss: 0.032561372965574265
step: 180, loss: 0.03747125715017319
step: 190, loss: 0.0020053619518876076
/usr/local/lib/python3.7/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=0.7098591549295775, f1=0.7713498622589531,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002025655936449766
step: 10, loss: 0.0009890692308545113
step: 20, loss: 0.03883179649710655
step: 30, loss: 0.0011080345138907433
step: 40, loss: 0.0009908447973430157
step: 50, loss: 0.0026590493507683277
step: 60, loss: 0.014310172758996487
step: 70, loss: 0.001585531048476696
step: 80, loss: 0.002290943404659629
step: 90, loss: 0.005725384224206209
step: 100, loss: 0.0007798545993864536
step: 110, loss: 0.0007579815573990345
step: 120, loss: 0.005091473460197449
step: 130, loss: 0.0006284798728302121
step: 140, loss: 0.003312833607196808
step: 150, loss: 0.0016235164366662502
step: 160, loss: 0.035467494279146194
step: 170, loss: 0.007355355191975832
step: 180, loss: 0.012537923641502857
step: 190, loss: 0.0007878179894760251
/usr/local/lib/python3.7/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=0.7159090909090909, f1=0.7648725212464589,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016024349024519324
step: 10, loss: 0.10520899295806885
step: 20, loss: 0.0039954474195837975
step: 30, loss: 0.0893360897898674
step: 40, loss: 0.004986220970749855
step: 50, loss: 0.016018696129322052
step: 60, loss: 0.0012389167677611113
step: 70, loss: 0.005464429967105389
step: 80, loss: 0.00106110330671072
step: 90, loss: 0.0016178929945454001
step: 100, loss: 0.0027328168507665396
step: 110, loss: 0.0002535840612836182
step: 120, loss: 0.0008657902362756431
step: 130, loss: 0.0006642909138463438
step: 140, loss: 0.07306279242038727
step: 150, loss: 0.007165221497416496
step: 160, loss: 0.04366856813430786
step: 170, loss: 0.0056288521736860275
step: 180, loss: 0.0005566072068177164
step: 190, loss: 0.05358187481760979
/usr/local/lib/python3.7/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=0.7113702623906706, f1=0.7572254335260116,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010980837978422642
step: 10, loss: 0.030665257945656776
step: 20, loss: 0.0014672112883999944
step: 30, loss: 0.00978055503219366
step: 40, loss: 0.0012156817829236388
step: 50, loss: 0.0038382408674806356
step: 60, loss: 0.0005931350169703364
step: 70, loss: 0.04086551070213318
step: 80, loss: 0.008511579595506191
step: 90, loss: 0.0007784241461195052
step: 100, loss: 0.0010023167124018073
step: 110, loss: 0.0007874926086515188
step: 120, loss: 0.0011793693993240595
step: 130, loss: 0.006900197826325893
step: 140, loss: 0.0028195541817694902
step: 150, loss: 0.002608741633594036
step: 160, loss: 0.1382601112127304
step: 170, loss: 0.004745216574519873
step: 180, loss: 0.01909925602376461

```

```

step: 190, loss: 0.010697593912482262
/usr/local/lib/python3.7/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=0.718918918918919, f1=0.7754010695187165,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022236083168536425
step: 10, loss: 0.0009659386123530567
step: 20, loss: 0.0032213947270065546
step: 30, loss: 0.0017639693105593324
step: 40, loss: 0.0013229313772171736
step: 50, loss: 0.0028563791420310736
step: 60, loss: 0.0009613787406124175
step: 70, loss: 0.003507094457745552
step: 80, loss: 0.00048432688345201313
step: 90, loss: 0.03174570947885513
step: 100, loss: 0.0019246068550273776
step: 110, loss: 0.0019258995307609439
step: 120, loss: 0.0011627598432824016
step: 130, loss: 0.0014188786735758185
step: 140, loss: 0.00029587611788883805
step: 150, loss: 0.009012987837195396
step: 160, loss: 0.01714254915714264
step: 170, loss: 0.029274752363562584
step: 180, loss: 0.005004190374165773
step: 190, loss: 0.0003179539635311812
/usr/local/lib/python3.7/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=0.7161803713527852, f1=0.769230769230769,
best_f1=0.7814207650273224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036704624071717262
step: 10, loss: 0.004819278139621019
step: 20, loss: 0.0318351686000824
step: 30, loss: 0.0004584514244925231
step: 40, loss: 0.000540051085408777
step: 50, loss: 0.0121309207752347
step: 60, loss: 0.0006679625366814435
step: 70, loss: 0.0006034759571775794
step: 80, loss: 0.0005803793319500983
step: 90, loss: 0.0030963029712438583
step: 100, loss: 0.001084402552805841

```



```

step: 110, loss: 0.0007600823300890625
step: 120, loss: 0.005655186716467142
step: 130, loss: 0.000992926536127925
step: 140, loss: 0.024170931428670883
step: 150, loss: 0.0007046050159260631
step: 160, loss: 0.0005298140458762646
step: 170, loss: 0.001283168327063322
step: 180, loss: 0.0008072001510299742
step: 190, loss: 0.008832739666104317
/usr/local/lib/python3.7/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=0.7172011661807581, f1=0.7586206896551724,
best_f1=0.7814207650273224
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 [00:06, 297.58it/s]
load_f1 = 0.6735218508997428
real_f1 = 0.656641604010025
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 [00:11, 370.91it/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 distilbert \
  --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 distilbert \
  --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: 0.8419774174690247
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.22708211839199066
step: 20, loss: 0.15217842161655426
step: 30, loss: 0.24339157342910767
step: 40, loss: 0.30987823009490967
step: 50, loss: 0.3750401437282562
step: 60, loss: 0.4385172724723816
step: 70, loss: 0.3164932429790497
step: 80, loss: 0.2523908317089081
step: 90, loss: 0.4077465236186981
step: 100, loss: 0.23591092228889465
step: 110, loss: 0.18716348707675934
step: 120, loss: 0.5620465278625488
step: 130, loss: 0.4143601059913635
step: 140, loss: 0.470867782831192

```

```

step: 150, loss: 0.11596797406673431
step: 160, loss: 0.34336531162261963
step: 170, loss: 0.22635814547538757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3643564356435644, f1=0.2916666666666667,
best_f1=0.2916666666666667
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3838541507720947
step: 10, loss: 0.17618486285209656
step: 20, loss: 0.2526737451553345
step: 30, loss: 0.2566850483417511
step: 40, loss: 0.19564348459243774
step: 50, loss: 0.1917034238576889
step: 60, loss: 0.15118777751922607
step: 70, loss: 0.22164559364318848
step: 80, loss: 0.1438918113708496
step: 90, loss: 0.25139039754867554
step: 100, loss: 0.10802258551120758
step: 110, loss: 0.20086923241615295
step: 120, loss: 0.12305604666471481
step: 130, loss: 0.15366309881210327
step: 140, loss: 0.26481664180755615
step: 150, loss: 0.08507510274648666
step: 160, loss: 0.09776116907596588
step: 170, loss: 0.2561298906803131
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6797385620915032, f1=0.6821052631578948,
best_f1=0.6821052631578948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08131836354732513
step: 10, loss: 0.08521129935979843
step: 20, loss: 0.08956887573003769

```

```

step: 30, loss: 0.30989980697631836
step: 40, loss: 0.010653706267476082
step: 50, loss: 0.10436760634183884
step: 60, loss: 0.16728821396827698
step: 70, loss: 0.09191334992647171
step: 80, loss: 0.1748117208480835
step: 90, loss: 0.12005694955587387
step: 100, loss: 0.03135054558515549
step: 110, loss: 0.08377783000469208
step: 120, loss: 0.005922188516706228
step: 130, loss: 0.3046424090862274
step: 140, loss: 0.014370723627507687
step: 150, loss: 0.07430266588926315
step: 160, loss: 0.047797493636608124
step: 170, loss: 0.21570225059986115
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7731958762886598, f1=0.7345971563981043,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05799998715519905
step: 10, loss: 0.13390524685382843
step: 20, loss: 0.02963419258594513
step: 30, loss: 0.09926989674568176
step: 40, loss: 0.011175074614584446
step: 50, loss: 0.07469592988491058
step: 60, loss: 0.08949850499629974
step: 70, loss: 0.038688648492097855
step: 80, loss: 0.05447342246770859
step: 90, loss: 0.034000467509031296
step: 100, loss: 0.04684001952409744
step: 110, loss: 0.13944639265537262
step: 120, loss: 0.14395570755004883
step: 130, loss: 0.05624699965119362
step: 140, loss: 0.03884583339095116
step: 150, loss: 0.018468597903847694
step: 160, loss: 0.0972215086221695
step: 170, loss: 0.011360126547515392
/usr/local/lib/python3.7/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=0.7584415584415585, f1=0.7179487179487181,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0355791300535202
step: 10, loss: 0.12976181507110596
step: 20, loss: 0.018642006441950798
step: 30, loss: 0.07268442213535309
step: 40, loss: 0.024815279990434647
step: 50, loss: 0.036213506013154984
step: 60, loss: 0.0029974260833114386
step: 70, loss: 0.03859136253595352
step: 80, loss: 0.0168579313904047
step: 90, loss: 0.053866539150476456
step: 100, loss: 0.0369185134768486
step: 110, loss: 0.20823366940021515
step: 120, loss: 0.02674885466694832
step: 130, loss: 0.009729673154652119
step: 140, loss: 0.032651130110025406
step: 150, loss: 0.00801536813378334
step: 160, loss: 0.06561911106109619
step: 170, loss: 0.12384937703609467
/usr/local/lib/python3.7/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=0.7574931880108993, f1=0.7543424317617866,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035528045147657394
step: 10, loss: 0.13025444746017456
step: 20, loss: 0.008442438207566738
step: 30, loss: 0.019673606380820274
step: 40, loss: 0.08625384420156479
step: 50, loss: 0.06279463320970535
step: 60, loss: 0.05136929079890251
step: 70, loss: 0.004358632955700159
step: 80, loss: 0.22630749642848969
step: 90, loss: 0.07433642446994781
step: 100, loss: 0.0727347880601883
step: 110, loss: 0.028462639078497887
step: 120, loss: 0.03517176955938339
step: 130, loss: 0.07746914774179459
step: 140, loss: 0.023815728724002838
step: 150, loss: 0.196994811296463
step: 160, loss: 0.1116378977894783

```

```

step: 170, loss: 0.03994583338499069
/usr/local/lib/python3.7/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=0.7297297297297297, f1=0.7764127764127765,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009500082582235336
step: 10, loss: 0.0029316144064068794
step: 20, loss: 0.028310511261224747
step: 30, loss: 0.004797115921974182
step: 40, loss: 0.04035334289073944
step: 50, loss: 0.00322687067091465
step: 60, loss: 0.26943865418434143
step: 70, loss: 0.06587788462638855
step: 80, loss: 0.004615491256117821
step: 90, loss: 0.052172642201185226
step: 100, loss: 0.001991876168176532
step: 110, loss: 0.057678695768117905
step: 120, loss: 0.09303313493728638
step: 130, loss: 0.4032585322856903
step: 140, loss: 0.02169387973845005
step: 150, loss: 0.11791221797466278
step: 160, loss: 0.07194844633340836
step: 170, loss: 0.16711239516735077
/usr/local/lib/python3.7/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=0.752577319587629, f1=0.7402298850574712,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018025238066911697
step: 10, loss: 0.006597120780497789
step: 20, loss: 0.00420360779389739
step: 30, loss: 0.025544384494423866
step: 40, loss: 0.03491666167974472
step: 50, loss: 0.01886880397796631
step: 60, loss: 0.009561481885612011
step: 70, loss: 0.03764282166957855
step: 80, loss: 0.01046591717749834
step: 90, loss: 0.03145476058125496
step: 100, loss: 0.053702693432569504
step: 110, loss: 0.01882406882941723
step: 120, loss: 0.015918290242552757

```

```

step: 130, loss: 0.0016102027148008347
step: 140, loss: 0.031069396063685417
step: 150, loss: 0.0965307280421257
step: 160, loss: 0.10187799483537674
step: 170, loss: 0.007393289357423782
/usr/local/lib/python3.7/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=0.7486631016042781, f1=0.745679012345679,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025893568992614746
step: 10, loss: 0.1519620716571808
step: 20, loss: 0.0045202807523310184
step: 30, loss: 0.011832332238554955
step: 40, loss: 0.06894797086715698
step: 50, loss: 0.007521153427660465
step: 60, loss: 0.3249688446521759
step: 70, loss: 0.05448165535926819
step: 80, loss: 0.07689402997493744
step: 90, loss: 0.017477624118328094
step: 100, loss: 0.023952478542923927
step: 110, loss: 0.004598662722855806
step: 120, loss: 0.06609869748353958
step: 130, loss: 0.0006944612832739949
step: 140, loss: 0.009560881182551384
step: 150, loss: 0.010351305827498436
step: 160, loss: 0.10749869048595428
step: 170, loss: 0.032650504261255264
/usr/local/lib/python3.7/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=0.7715736040609138, f1=0.7523364485981308,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005838570650666952
step: 10, loss: 0.023353369906544685
step: 20, loss: 0.003985259216278791
step: 30, loss: 0.003015994792804122
step: 40, loss: 0.009755131788551807
step: 50, loss: 0.015443711541593075
step: 60, loss: 0.002160724950954318
step: 70, loss: 0.013075160793960094
step: 80, loss: 0.16406016051769257

```



```

step: 90, loss: 0.018594278022646904
step: 100, loss: 0.06113388016819954
step: 110, loss: 0.006071661598980427
step: 120, loss: 0.043817538768053055
step: 130, loss: 0.19237428903579712
step: 140, loss: 0.07601822167634964
step: 150, loss: 0.0055517894215881824
step: 160, loss: 0.011612235568463802
step: 170, loss: 0.017295707017183304
/usr/local/lib/python3.7/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=0.7480106100795756, f1=0.7524271844660194,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007467292598448694
step: 10, loss: 0.020577700808644295
step: 20, loss: 0.08686736971139908
step: 30, loss: 0.008480586111545563
step: 40, loss: 0.0604533888399601
step: 50, loss: 0.0020748525857925415
step: 60, loss: 0.01674177125096321
step: 70, loss: 0.00798093806952238
step: 80, loss: 0.0017260911408811808
step: 90, loss: 0.005747444927692413
step: 100, loss: 0.0008847851422615349
step: 110, loss: 0.0010076535400003195
step: 120, loss: 0.11666920781135559
step: 130, loss: 0.008787288330495358
step: 140, loss: 0.054353900253772736
step: 150, loss: 0.017562801018357277
step: 160, loss: 0.0033088696654886007
step: 170, loss: 0.07039463520050049
/usr/local/lib/python3.7/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=0.7417582417582418, f1=0.7411167512690355,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008909433148801327
step: 10, loss: 0.019846096634864807
step: 20, loss: 0.022816479206085205
step: 30, loss: 0.0010309041244909167
step: 40, loss: 0.027627743780612946

```

```

step: 50, loss: 0.0060383533127605915
step: 60, loss: 0.08761300891637802
step: 70, loss: 0.04487496241927147
step: 80, loss: 0.15595850348472595
step: 90, loss: 0.0008023007540032268
step: 100, loss: 0.0018156690057367086
step: 110, loss: 0.001700012944638729
step: 120, loss: 0.014552361331880093
step: 130, loss: 0.01200181059539318
step: 140, loss: 0.0006746642175130546
step: 150, loss: 0.009250383824110031
step: 160, loss: 0.0021150833927094936
step: 170, loss: 0.0032326264772564173
/usr/local/lib/python3.7/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=0.7533875338753389, f1=0.7341772151898734,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005263368599116802
step: 10, loss: 0.011883457191288471
step: 20, loss: 0.0014566902536898851
step: 30, loss: 0.005842709913849831
step: 40, loss: 0.004073336720466614
step: 50, loss: 0.0011260820319876075
step: 60, loss: 0.00435453187674284
step: 70, loss: 0.04788028821349144
step: 80, loss: 0.004840930923819542
step: 90, loss: 0.0009600009652785957
step: 100, loss: 0.000555488804820925
step: 110, loss: 0.008884130045771599
step: 120, loss: 0.003054654225707054
step: 130, loss: 0.01317108515650034
step: 140, loss: 0.0018933262908831239
step: 150, loss: 0.002734130946919322
step: 160, loss: 0.001274880487471819
step: 170, loss: 0.003762230509892106
/usr/local/lib/python3.7/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=0.7596899224806202, f1=0.7445255474452555,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014684398658573627

```

```

step: 10, loss: 0.0020139687694609165
step: 20, loss: 0.023045193403959274
step: 30, loss: 0.00733172008767724
step: 40, loss: 0.0014893451007083058
step: 50, loss: 0.19585546851158142
step: 60, loss: 0.0021272865124046803
step: 70, loss: 0.027847060933709145
step: 80, loss: 0.0009094639681279659
step: 90, loss: 0.003798371646553278
step: 100, loss: 0.01205003634095192
step: 110, loss: 0.0018051554216071963
step: 120, loss: 0.0018289305735379457
step: 130, loss: 0.014937723986804485
step: 140, loss: 0.00657698605209589
step: 150, loss: 0.002690758090466261
step: 160, loss: 0.0282841045409441
step: 170, loss: 0.0038495443295687437
/usr/local/lib/python3.7/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=0.7642276422764228, f1=0.7461928934010152,
best_f1=0.7345971563981043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01259460486471653
step: 10, loss: 0.003586304374039173
step: 20, loss: 0.08323115110397339
step: 30, loss: 0.010906081646680832
step: 40, loss: 0.0006485236808657646
step: 50, loss: 0.0056292153894901276
step: 60, loss: 0.0009966279612854123
step: 70, loss: 0.021656977012753487
step: 80, loss: 0.0026747863739728928
step: 90, loss: 0.00038917968049645424
step: 100, loss: 0.0027009695768356323
step: 110, loss: 0.0006117206648923457
step: 120, loss: 0.026120228692889214
step: 130, loss: 0.005783436354249716
step: 140, loss: 0.002365215914323926
step: 150, loss: 0.006647906266152859
step: 160, loss: 0.000492981867864728
step: 170, loss: 0.0026228020433336496
/usr/local/lib/python3.7/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=0.7580645161290323, f1=0.7412935323383084,
best_f1=0.7345971563981043

```

```
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 [00:04, 430.43it/s]
load_f1 = 0.5258215962441315
real_f1 = 0.4945295404814004
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 [00:11, 368.87it/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 distilbert \
--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 distilbert \
--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: 0.8085821866989136
step: 10, loss: 0.47838881611824036
step: 20, loss: 0.5703800916671753
step: 30, loss: 0.4836879372596741
step: 40, loss: 0.32405537366867065
step: 50, loss: 0.2353421300649643
step: 60, loss: 0.13976368308067322
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.1447216421365738
step: 80, loss: 0.25415581464767456
step: 90, loss: 0.11759679764509201
step: 100, loss: 0.18342703580856323
step: 110, loss: 0.08494134247303009
step: 120, loss: 0.017019622027873993
step: 130, loss: 0.02678391896188259
step: 140, loss: 0.09186301380395889
step: 150, loss: 0.1818416714668274
step: 160, loss: 0.2871900796890259
step: 170, loss: 0.018716994673013687
step: 180, loss: 0.006356722675263882
step: 190, loss: 0.07880502939224243
step: 200, loss: 0.05989193543791771
step: 210, loss: 0.03337070718407631
step: 220, loss: 0.030812636017799377
step: 230, loss: 0.11894208192825317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9535398230088497, f1=0.9469026548672567,
best_f1=0.9469026548672567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012298576533794403
step: 10, loss: 0.003225144697353244
step: 20, loss: 0.016420476138591766
step: 30, loss: 0.019013866782188416
step: 40, loss: 0.0320497527718544
step: 50, loss: 0.04247823357582092
step: 60, loss: 0.004565211478620768
step: 70, loss: 0.008311552926898003
step: 80, loss: 0.010670213960111141
step: 90, loss: 0.004548025317490101
step: 100, loss: 0.2541874051094055
step: 110, loss: 0.06256464123725891
step: 120, loss: 0.14879131317138672
step: 130, loss: 0.021781733259558678
step: 140, loss: 0.1714954674243927
step: 150, loss: 0.12749744951725006
step: 160, loss: 0.012494975700974464
step: 170, loss: 0.09553302824497223
step: 180, loss: 0.014228198677301407
step: 190, loss: 0.08233635127544403
step: 200, loss: 0.05194198712706566
step: 210, loss: 0.09817610681056976
step: 220, loss: 0.0012782923877239227
step: 230, loss: 0.03429458662867546
/usr/local/lib/python3.7/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=0.9534368070953436, f1=0.9654403567447045,
best_f1=0.9469026548672567
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08925629407167435
step: 10, loss: 0.023302221670746803
step: 20, loss: 0.005389802623540163
step: 30, loss: 0.09004190564155579
step: 40, loss: 0.09259786456823349
step: 50, loss: 0.025584090501070023
step: 60, loss: 0.07014194130897522
step: 70, loss: 0.0011113624786958098
```

```

step: 80, loss: 0.025634344667196274
step: 90, loss: 0.17460788786411285
step: 100, loss: 0.005188563372939825
step: 110, loss: 0.054726146161556244
step: 120, loss: 0.023062655702233315
step: 130, loss: 0.0015987895894795656
step: 140, loss: 0.0071175131015479565
step: 150, loss: 0.004298201762139797
step: 160, loss: 0.004960104823112488
step: 170, loss: 0.00741219986230135
step: 180, loss: 0.006102230399847031
step: 190, loss: 0.026602761819958687
step: 200, loss: 0.017455261200666428
step: 210, loss: 0.011120990850031376
step: 220, loss: 0.08513013273477554
step: 230, loss: 0.05935061350464821
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9575892857142857, f1=0.9450056116722784,
best_f1=0.9450056116722784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023260097950696945
step: 10, loss: 0.016091344878077507
step: 20, loss: 0.011033373884856701
step: 30, loss: 0.006554227322340012
step: 40, loss: 0.018321320414543152
step: 50, loss: 0.0020865192636847496
step: 60, loss: 0.01282312162220478
step: 70, loss: 0.02529081515967846
step: 80, loss: 0.12256431579589844
step: 90, loss: 0.01900079846382141
step: 100, loss: 0.004306082613766193
step: 110, loss: 0.030672205612063408
step: 120, loss: 0.0963188037276268
step: 130, loss: 0.053183846175670624
step: 140, loss: 0.010827787220478058
step: 150, loss: 0.023198671638965607
step: 160, loss: 0.07701738178730011
step: 170, loss: 0.009202595800161362
step: 180, loss: 0.10673914104700089
step: 190, loss: 0.05952581763267517

```



```

step: 200, loss: 0.001315073692239821
step: 210, loss: 0.018742695450782776
step: 220, loss: 0.003586419392377138
step: 230, loss: 0.00038533148472197354
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9601769911504424, f1=0.9643652561247216,
best_f1=0.9643652561247216
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015230280114337802
step: 10, loss: 0.0013013763818889856
step: 20, loss: 0.0006798499962314963
step: 30, loss: 0.0006026571500115097
step: 40, loss: 0.00048020173562690616
step: 50, loss: 0.001034334534779191
step: 60, loss: 0.0036778717767447233
step: 70, loss: 0.010953173041343689
step: 80, loss: 0.011081437580287457
step: 90, loss: 0.0054132211953401566
step: 100, loss: 0.0023959632962942123
step: 110, loss: 0.001033485634252429
step: 120, loss: 0.05629583075642586
step: 130, loss: 0.0339767225086689
step: 140, loss: 0.0002737016184255481
step: 150, loss: 0.0003288665320724249
step: 160, loss: 0.15011724829673767
step: 170, loss: 0.1571034938097
step: 180, loss: 0.007560944650322199
step: 190, loss: 0.004160861950367689
step: 200, loss: 0.0004905643290840089
step: 210, loss: 0.00044252301449887455
step: 220, loss: 0.0011539030820131302
step: 230, loss: 0.0010789962252601981
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9646017699115044, f1=0.9591160220994475,

```

```

best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003969492681790143
step: 10, loss: 0.005170351825654507
step: 20, loss: 0.001623230054974556
step: 30, loss: 0.0002191313833463937
step: 40, loss: 0.007596069481223822
step: 50, loss: 0.021507656201720238
step: 60, loss: 0.0017521060071885586
step: 70, loss: 0.0008198090363293886
step: 80, loss: 0.006646201014518738
step: 90, loss: 0.0008708559907972813
step: 100, loss: 0.0003578817704692483
step: 110, loss: 0.023478858172893524
step: 120, loss: 0.00023074982163961977
step: 130, loss: 0.00030027766479179263
step: 140, loss: 0.043486081063747406
step: 150, loss: 0.006942518521100283
step: 160, loss: 0.0711209625005722
step: 170, loss: 0.002856291364878416
step: 180, loss: 0.0026610642671585083
step: 190, loss: 0.0028154875617474318
step: 200, loss: 0.03579030558466911
step: 210, loss: 0.003062634263187647
step: 220, loss: 0.0013487177202478051
step: 230, loss: 0.0003035056288354099
/usr/local/lib/python3.7/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=0.96353591160221, f1=0.9591160220994475,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04231051355600357
step: 10, loss: 0.006975471507757902
step: 20, loss: 0.0003643813543021679
step: 30, loss: 0.0017884044209495187
step: 40, loss: 0.0011466366704553366
step: 50, loss: 0.00010261948045808822
step: 60, loss: 0.008214316330850124
step: 70, loss: 0.0021747334394603968
step: 80, loss: 0.00014706538058817387
step: 90, loss: 0.030861718580126762
step: 100, loss: 0.0031375624239444733
step: 110, loss: 0.003503964049741626

```

```

step: 120, loss: 0.011346565559506416
step: 130, loss: 0.0004459054907783866
step: 140, loss: 0.00013007612142246217
step: 150, loss: 0.0023734692949801683
step: 160, loss: 0.00021233038569334894
step: 170, loss: 0.00016998428327497095
step: 180, loss: 0.00018035103857982904
step: 190, loss: 0.0005808814312331378
step: 200, loss: 0.22587324678897858
step: 210, loss: 0.006716250441968441
step: 220, loss: 0.001170082134194672
step: 230, loss: 0.006371622905135155
/usr/local/lib/python3.7/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=0.9636163175303197, f1=0.9633740288568259,
best_f1=0.9591160220994475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04472828283905983
step: 10, loss: 0.001716800034046173
step: 20, loss: 0.0001760418963385746
step: 30, loss: 0.0071113938465714455
step: 40, loss: 0.0008947554742917418
step: 50, loss: 0.0006852173246443272
step: 60, loss: 0.00017728043894749135
step: 70, loss: 0.0002172076638089493
step: 80, loss: 0.0005777691840194166
step: 90, loss: 0.00043507415102794766
step: 100, loss: 0.0012102953623980284
step: 110, loss: 0.021907540038228035
step: 120, loss: 0.00017694821872282773
step: 130, loss: 0.09016890078783035
step: 140, loss: 0.00019735380192287266
step: 150, loss: 0.00024673217558301985
step: 160, loss: 0.0027931102085858583
step: 170, loss: 0.00021036979160271585
step: 180, loss: 0.008636927232146263
step: 190, loss: 0.00013183712144382298
step: 200, loss: 0.03628282994031906
step: 210, loss: 0.00043282724800519645
step: 220, loss: 0.0004125208070036024
step: 230, loss: 0.03587713837623596
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9675977653631285, f1=0.9675977653631285,

best_f1=0.9675977653631285

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

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

step: 0, loss: 0.0011568247573450208

step: 10, loss: 0.0002836317871697247

step: 20, loss: 0.013983168639242649

step: 30, loss: 0.004565212409943342

step: 40, loss: 0.0007100813672877848

step: 50, loss: 0.0001787236542440951

step: 60, loss: 0.0011458487715572119

step: 70, loss: 0.001585057470947504

step: 80, loss: 0.03393996134400368

step: 90, loss: 0.0012420571874827147

step: 100, loss: 0.00016409643285442144

step: 110, loss: 7.559865480288863e-05

step: 120, loss: 0.009513158351182938

step: 130, loss: 8.031162724364549e-05

step: 140, loss: 0.013663174584507942

step: 150, loss: 7.378185546258464e-05

step: 160, loss: 0.00020276676514185965

step: 170, loss: 0.000267460651230067

step: 180, loss: 0.006420222111046314

step: 190, loss: 0.00028919894248247147

step: 200, loss: 0.0010007653618231416

step: 210, loss: 0.0002494606887921691

step: 220, loss: 0.027452921494841576

step: 230, loss: 0.009829333983361721

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

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

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

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=0.972129319955407, f1=0.9644444444444443,

best_f1=0.9644444444444443

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

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

step: 0, loss: 0.00017364499217364937

step: 10, loss: 5.226311623118818e-05

step: 20, loss: 0.0012879881542176008

step: 30, loss: 0.0001694071979727596

```

step: 40, loss: 0.002103667240589857
step: 50, loss: 0.11378820985555649
step: 60, loss: 0.04296226426959038
step: 70, loss: 0.00016668302123434842
step: 80, loss: 0.0006275305058807135
step: 90, loss: 0.00018672921578399837
step: 100, loss: 0.00016492231225129217
step: 110, loss: 0.0005252960836514831
step: 120, loss: 0.01561708189547062
step: 130, loss: 0.008396238088607788
step: 140, loss: 0.03625616058707237
step: 150, loss: 0.00025240343529731035
step: 160, loss: 0.00019939996127504855
step: 170, loss: 0.004100382793694735
step: 180, loss: 0.0004008722025901079
step: 190, loss: 0.11434485018253326
step: 200, loss: 0.001739738043397665
step: 210, loss: 0.00012731837341561913
step: 220, loss: 0.005612798035144806
step: 230, loss: 0.010418335907161236
/usr/local/lib/python3.7/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=0.9707865168539327, f1=0.9673790776152981,
best_f1=0.9644444444444443
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003034706460312009
step: 10, loss: 0.00029970044852234423
step: 20, loss: 0.00011858416837640107
step: 30, loss: 0.016813507303595543
step: 40, loss: 0.006430634297430515
step: 50, loss: 0.0012311800383031368
step: 60, loss: 0.0003645865654107183
step: 70, loss: 9.500111627858132e-05
step: 80, loss: 0.0002098301483783871
step: 90, loss: 0.00044865295058116317
step: 100, loss: 0.00012470850197132677
step: 110, loss: 8.5858438978903e-05
step: 120, loss: 0.0029420999344438314
step: 130, loss: 0.00014590135833714157
step: 140, loss: 0.0019738811533898115
step: 150, loss: 4.972556052962318e-05
step: 160, loss: 8.308492397191003e-05
step: 170, loss: 0.019529204815626144
step: 180, loss: 0.03060128539800644
step: 190, loss: 0.006339447572827339

```

```

step: 200, loss: 0.0027008901815861464
step: 210, loss: 6.293001933954656e-05
step: 220, loss: 0.00019037675519939512
step: 230, loss: 0.015583890490233898
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9732739420935412, f1=0.9642058165548099,
best_f1=0.9642058165548099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007452602731063962
step: 10, loss: 0.004274706821888685
step: 20, loss: 6.494112312793732e-05
step: 30, loss: 0.00038889888674020767
step: 40, loss: 0.0003009962965734303
step: 50, loss: 0.00012398338003549725
step: 60, loss: 0.01336623728275299
step: 70, loss: 0.0005223980406299233
step: 80, loss: 0.0009594119037501514
step: 90, loss: 0.002192866522818804
step: 100, loss: 6.286564166657627e-05
step: 110, loss: 0.0011697950540110469
step: 120, loss: 0.0002599572471808642
step: 130, loss: 0.351036936044693
step: 140, loss: 0.000739346956834197
step: 150, loss: 0.0005211070529185236
step: 160, loss: 0.01276439055800438
step: 170, loss: 8.514027285855263e-05
step: 180, loss: 0.0010426660301163793
step: 190, loss: 0.01632305607199669
step: 200, loss: 0.017519574612379074
step: 210, loss: 0.0005234418204054236
step: 220, loss: 0.02316601574420929
step: 230, loss: 0.00012913237151224166
/usr/local/lib/python3.7/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=0.9721913236929923, f1=0.9621380846325166,
best_f1=0.9642058165548099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0002135664690285921
step: 10, loss: 0.0001784776250133291
step: 20, loss: 7.416383596137166e-05
step: 30, loss: 0.02471444383263588
step: 40, loss: 0.00043811893556267023
step: 50, loss: 0.00010721573198679835
step: 60, loss: 0.0013530259020626545
step: 70, loss: 4.817304579773918e-05
step: 80, loss: 0.0001717151899356395
step: 90, loss: 6.898290303070098e-05
step: 100, loss: 6.967041554162279e-05
step: 110, loss: 0.0015660665230825543
step: 120, loss: 0.00035874336026608944
step: 130, loss: 0.0010577649809420109
step: 140, loss: 0.000697324110660702
step: 150, loss: 0.0035418469924479723
step: 160, loss: 5.600273289019242e-05
step: 170, loss: 0.000267225899733603
step: 180, loss: 0.00012955872807651758
step: 190, loss: 0.030101057142019272
step: 200, loss: 0.00012022642476949841
step: 210, loss: 0.00011090078623965383
step: 220, loss: 0.0004709255590569228
step: 230, loss: 7.437846943503246e-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=0.9733924611973392, f1=0.9655937846836848,
best_f1=0.9655937846836848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002250795077998191
step: 10, loss: 3.94384769606404e-05
step: 20, loss: 0.01909412257373333
step: 30, loss: 0.001956619555130601
step: 40, loss: 5.182397217140533e-05
step: 50, loss: 0.0009435269166715443
step: 60, loss: 0.004277147352695465
step: 70, loss: 5.93091499467846e-05
step: 80, loss: 0.0001266846084035933
step: 90, loss: 0.00014038383960723877
step: 100, loss: 0.005018720403313637
step: 110, loss: 0.013871412724256516

```

```

step: 120, loss: 0.0001495737669756636
step: 130, loss: 0.00010726073378464207
step: 140, loss: 0.0038645125459879637
step: 150, loss: 0.0002408657019259408
step: 160, loss: 7.244189328048378e-05
step: 170, loss: 5.9559362853178754e-05
step: 180, loss: 0.0002449331514071673
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.07668064534664154
step: 200, loss: 0.00027800616226159036
step: 210, loss: 0.00014759867917746305
step: 220, loss: 6.836318789282814e-05
step: 230, loss: 0.00013903758372180164
/usr/local/lib/python3.7/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=0.9733333333333333, f1=0.9644444444444443,
best_f1=0.9655937846836848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.1654144044732675e-05
step: 10, loss: 0.0004727831110358238
step: 20, loss: 8.440860983682796e-05
step: 30, loss: 0.0032464549876749516
step: 40, loss: 9.37722870730795e-05
step: 50, loss: 7.657490641577169e-05
step: 60, loss: 4.499153874348849e-05
step: 70, loss: 6.378255784511566e-05
step: 80, loss: 0.0005696197622455657
step: 90, loss: 4.054775126860477e-05
step: 100, loss: 9.106638026423752e-05
step: 110, loss: 3.9691840356681496e-05
step: 120, loss: 0.14540696144104004
step: 130, loss: 0.00012110213720006868
step: 140, loss: 0.00030426119337789714
step: 150, loss: 0.00020145390590187162
step: 160, loss: 0.01600831001996994
step: 170, loss: 3.1526484235655516e-05
step: 180, loss: 0.00041374596185050905
step: 190, loss: 9.563133062329143e-05
step: 200, loss: 5.996132313157432e-05
step: 210, loss: 0.0042937095277011395
step: 220, loss: 0.0027313160244375467
step: 230, loss: 4.4918004277860746e-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=0.9733333333333333, f1=0.9644444444444443,
best_f1=0.9655937846836848
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.")
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 [00:07, 316.50it/s]
load_f1 = 0.9720670391061451
real_f1 = 0.9656699889258028
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 [00:11, 371.41it/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 distilbert \
--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 distilbert \
--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: 0.7942676544189453
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44365087151527405
step: 20, loss: 0.5118536949157715
step: 30, loss: 0.45975354313850403
step: 40, loss: 0.45389556884765625
step: 50, loss: 0.3982391059398651
step: 60, loss: 0.37856611609458923
step: 70, loss: 0.1529812067747116
step: 80, loss: 0.25887390971183777
step: 90, loss: 0.21557557582855225
step: 100, loss: 0.24549925327301025
step: 110, loss: 0.1600096970796585
step: 120, loss: 0.15005062520503998
step: 130, loss: 0.04675512760877609
step: 140, loss: 0.16512688994407654
step: 150, loss: 0.0633515939116478
step: 160, loss: 0.13164083659648895
step: 170, loss: 0.39179670810699463
step: 180, loss: 0.14735472202301025
step: 190, loss: 0.06138845905661583
step: 200, loss: 0.10543425381183624
step: 210, loss: 0.11100563406944275
step: 220, loss: 0.19622080028057098
step: 230, loss: 0.12837621569633484
step: 240, loss: 0.10263589769601822
step: 250, loss: 0.07807612419128418
step: 260, loss: 0.03650282695889473
step: 270, loss: 0.05805729702115059

```

```

step: 280, loss: 0.1663917750120163
step: 290, loss: 0.07734494656324387
step: 300, loss: 0.2363283336162567
step: 310, loss: 0.07297191768884659
step: 320, loss: 0.14061324298381805
step: 330, loss: 0.1798301637172699
step: 340, loss: 0.23575876653194427
step: 350, loss: 0.08089148998260498
step: 360, loss: 0.1226823702454567
step: 370, loss: 0.13269591331481934
step: 380, loss: 0.2493114322423935
step: 390, loss: 0.020507987588644028
step: 400, loss: 0.0342998169362545
step: 410, loss: 0.07963736355304718
step: 420, loss: 0.03539406880736351
step: 430, loss: 0.08002913743257523
step: 440, loss: 0.1991264373064041
step: 450, loss: 0.06946483254432678
step: 460, loss: 0.018291369080543518
step: 470, loss: 0.40249788761138916
step: 480, loss: 0.25246232748031616
step: 490, loss: 0.20062777400016785
step: 500, loss: 0.08891046792268753
step: 510, loss: 0.15879444777965546
step: 520, loss: 0.08881936222314835
step: 530, loss: 0.07956232875585556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.908411214953271, f1=0.9071494893221911,
best_f1=0.9071494893221911
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05946626141667366
step: 10, loss: 0.16491344571113586
step: 20, loss: 0.09410328418016434
step: 30, loss: 0.05331973731517792
step: 40, loss: 0.00895713921636343
step: 50, loss: 0.14067135751247406
step: 60, loss: 0.18810372054576874
step: 70, loss: 0.21829308569431305
step: 80, loss: 0.031106922775506973
step: 90, loss: 0.04518063738942146

```

```
step: 100, loss: 0.19078510999679565
step: 110, loss: 0.04246804490685463
step: 120, loss: 0.030147159472107887
step: 130, loss: 0.026474442332983017
step: 140, loss: 0.04386165738105774
step: 150, loss: 0.05630536004900932
step: 160, loss: 0.07673726230859756
step: 170, loss: 0.10708077996969223
step: 180, loss: 0.05419887602329254
step: 190, loss: 0.09373430907726288
step: 200, loss: 0.016652945429086685
step: 210, loss: 0.07356598228216171
step: 220, loss: 0.21749839186668396
step: 230, loss: 0.09195264428853989
step: 240, loss: 0.16193713247776031
step: 250, loss: 0.0963338315486908
step: 260, loss: 0.040448084473609924
step: 270, loss: 0.021068939939141273
step: 280, loss: 0.2241748720407486
step: 290, loss: 0.11264532059431076
step: 300, loss: 0.05023255944252014
step: 310, loss: 0.1596110463142395
step: 320, loss: 0.19277657568454742
step: 330, loss: 0.07261404395103455
step: 340, loss: 0.00836690329015255
step: 350, loss: 0.05200406536459923
step: 360, loss: 0.0487535335123539
step: 370, loss: 0.009028435684740543
step: 380, loss: 0.09038300812244415
step: 390, loss: 0.03562421724200249
step: 400, loss: 0.19891777634620667
step: 410, loss: 0.03744899481534958
step: 420, loss: 0.09445646405220032
step: 430, loss: 0.023826032876968384
step: 440, loss: 0.011067178100347519
step: 450, loss: 0.046542711555957794
step: 460, loss: 0.19235914945602417
step: 470, loss: 0.09117327630519867
step: 480, loss: 0.30056294798851013
step: 490, loss: 0.07862913608551025
step: 500, loss: 0.06680109351873398
step: 510, loss: 0.11088372766971588
step: 520, loss: 0.011571191251277924
step: 530, loss: 0.10490628331899643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9213587715216379, f1=0.9121996303142329,

best_f1=0.9121996303142329

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

An input tensor was not cuda.

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

step: 0, loss: 0.020259886980056763

step: 10, loss: 0.09672517329454422

step: 20, loss: 0.20173345506191254

step: 30, loss: 0.20056293904781342

step: 40, loss: 0.09775891155004501

step: 50, loss: 0.10821801424026489

step: 60, loss: 0.011876685544848442

step: 70, loss: 0.017219891771674156

step: 80, loss: 0.07334520667791367

step: 90, loss: 0.21531477570533752

step: 100, loss: 0.1484495848417282

step: 110, loss: 0.039206862449645996

step: 120, loss: 0.11459697037935257

step: 130, loss: 0.030040554702281952

step: 140, loss: 0.01764640584588051

step: 150, loss: 0.020841002464294434

step: 160, loss: 0.029580648988485336

step: 170, loss: 0.10631857067346573

step: 180, loss: 0.03353195637464523

step: 190, loss: 0.009304950945079327

step: 200, loss: 0.10393287986516953

step: 210, loss: 0.11450012028217316

step: 220, loss: 0.02900584600865841

step: 230, loss: 0.0252201110124588

step: 240, loss: 0.025178296491503716

step: 250, loss: 0.009081361815333366

step: 260, loss: 0.01496891863644123

step: 270, loss: 0.01660672016441822

step: 280, loss: 0.025556931272149086

step: 290, loss: 0.0731712281703949

step: 300, loss: 0.11281566321849823

step: 310, loss: 0.13755947351455688

step: 320, loss: 0.34630370140075684

step: 330, loss: 0.0037537449970841408

step: 340, loss: 0.005446307361125946

step: 350, loss: 0.08607736974954605

step: 360, loss: 0.019931767135858536

step: 370, loss: 0.012845207937061787

step: 380, loss: 0.04641193524003029

step: 390, loss: 0.012091044336557388

```

step: 400, loss: 0.0402589812874794
step: 410, loss: 0.02838161028921604
step: 420, loss: 0.03046688251197338
step: 430, loss: 0.03578044846653938
step: 440, loss: 0.09508626163005829
step: 450, loss: 0.1267554759979248
step: 460, loss: 0.1791960746049881
step: 470, loss: 0.02561964839696884
step: 480, loss: 0.015897415578365326
step: 490, loss: 0.09700490534305573
step: 500, loss: 0.09032152593135834
step: 510, loss: 0.013629257678985596
step: 520, loss: 0.008090736344456673
step: 530, loss: 0.02083488367497921
/usr/local/lib/python3.7/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=0.9162377164248947, f1=0.9139382600561271,
best_f1=0.9121996303142329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007583294063806534
step: 10, loss: 0.0157733503729105
step: 20, loss: 0.014650448225438595
step: 30, loss: 0.005322040989995003
step: 40, loss: 0.032758038491010666
step: 50, loss: 0.03164854273200035
step: 60, loss: 0.0018846420571208
step: 70, loss: 0.001279704156331718
step: 80, loss: 0.028225267305970192
step: 90, loss: 0.01263098418712616
step: 100, loss: 0.02203529141843319
step: 110, loss: 0.019387826323509216
step: 120, loss: 0.0007882615900598466
step: 130, loss: 0.09271437674760818
step: 140, loss: 0.09813850373029709
step: 150, loss: 0.014284245669841766
step: 160, loss: 0.02220410294830799
step: 170, loss: 0.06167340651154518
step: 180, loss: 0.04729454219341278
step: 190, loss: 0.023444654420018196
step: 200, loss: 0.004321706481277943
step: 210, loss: 0.16025961935520172
step: 220, loss: 0.013801334425807
step: 230, loss: 0.2120748609304428
step: 240, loss: 0.08473662286996841
step: 250, loss: 0.00913647748529911

```

```

step: 260, loss: 0.052923619747161865
step: 270, loss: 0.026587512344121933
step: 280, loss: 0.009401810355484486
step: 290, loss: 0.0053249867632985115
step: 300, loss: 0.0018189185066148639
step: 310, loss: 0.011984777636826038
step: 320, loss: 0.03166978806257248
step: 330, loss: 0.14742574095726013
step: 340, loss: 0.19022873044013977
step: 350, loss: 0.04235413298010826
step: 360, loss: 0.11399274319410324
step: 370, loss: 0.03064029850065708
step: 380, loss: 0.02580096945166588
step: 390, loss: 0.1516178697347641
step: 400, loss: 0.05100010335445404
step: 410, loss: 0.03489692509174347
step: 420, loss: 0.024844301864504814
step: 430, loss: 0.07149364054203033
step: 440, loss: 0.10407355427742004
step: 450, loss: 0.01213037595152855
step: 460, loss: 0.0036615924909710884
step: 470, loss: 0.026320843026041985
step: 480, loss: 0.008651413954794407
step: 490, loss: 0.08950722217559814
step: 500, loss: 0.02253641001880169
step: 510, loss: 0.06544704735279083
step: 520, loss: 0.13401800394058228
step: 530, loss: 0.003727729432284832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9228650137741048, f1=0.9174646602827177,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03604031726717949
step: 10, loss: 0.047457389533519745
step: 20, loss: 0.017711300402879715
step: 30, loss: 0.030493292957544327
step: 40, loss: 0.06770691275596619
step: 50, loss: 0.006410396192222834
step: 60, loss: 0.11388114094734192
step: 70, loss: 0.03517073392868042

```


step: 80, loss: 0.002151195891201496
step: 90, loss: 0.021503573283553123
step: 100, loss: 0.049356088042259216
step: 110, loss: 0.003651948180049658
step: 120, loss: 0.005848876666277647
step: 130, loss: 0.004039232153445482
step: 140, loss: 0.004116039723157883
step: 150, loss: 0.004972924944013357
step: 160, loss: 0.01268045138567686
step: 170, loss: 0.024443916976451874
step: 180, loss: 0.008264991454780102
step: 190, loss: 0.0016931085847318172
step: 200, loss: 0.01960541307926178
step: 210, loss: 0.0076171960681676865
step: 220, loss: 0.0004924268578179181
step: 230, loss: 0.0005644368939101696
step: 240, loss: 0.02494463324546814
step: 250, loss: 0.0004849615797866136
step: 260, loss: 0.04321427270770073
step: 270, loss: 0.09733542799949646
step: 280, loss: 0.02231323905289173
step: 290, loss: 0.0009855660609900951
step: 300, loss: 0.1087094396352768
step: 310, loss: 0.1851627677679062
step: 320, loss: 0.015261407941579819
step: 330, loss: 0.10633885860443115
step: 340, loss: 0.06524171680212021
step: 350, loss: 0.012396072968840599
step: 360, loss: 0.0044496869668364525
step: 370, loss: 0.001227743923664093
step: 380, loss: 0.0023968841414898634
step: 390, loss: 0.031073415651917458
step: 400, loss: 0.009137177839875221
step: 410, loss: 0.003614722518250346
step: 420, loss: 0.015150152146816254
step: 430, loss: 0.023763101547956467
step: 440, loss: 0.02669556997716427
step: 450, loss: 0.015942830592393875
step: 460, loss: 0.0452667661011219
step: 470, loss: 0.00997746642678976
step: 480, loss: 0.03636803850531578
step: 490, loss: 0.045806434005498886
step: 500, loss: 0.01522468589246273
step: 510, loss: 0.030337560921907425
step: 520, loss: 0.004739142023026943
step: 530, loss: 0.02747921645641327
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9249884951679707, f1=0.9225126088950022,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003339191898703575
step: 10, loss: 0.02243892475962639
step: 20, loss: 0.00040136263123713434
step: 30, loss: 0.0004525081894826144
step: 40, loss: 0.00804411806166172
step: 50, loss: 0.001964631024748087
step: 60, loss: 0.0023259101435542107
step: 70, loss: 0.029553717002272606
step: 80, loss: 0.033987388014793396
step: 90, loss: 0.0388147346675396
step: 100, loss: 0.030194099992513657
step: 110, loss: 0.005720177199691534
step: 120, loss: 0.003739973297342658
step: 130, loss: 0.01953510008752346
step: 140, loss: 0.0020418320782482624
step: 150, loss: 0.0059360540471971035
step: 160, loss: 0.0005258264718577266
step: 170, loss: 0.001231888309121132
step: 180, loss: 0.0011223176261410117
step: 190, loss: 0.02188444696366787
step: 200, loss: 0.001998536055907607
step: 210, loss: 0.011747895739972591
step: 220, loss: 0.006892107427120209
step: 230, loss: 0.0029610421042889357
step: 240, loss: 0.0009533856064081192
step: 250, loss: 0.002640216378495097
step: 260, loss: 0.007022282108664513
step: 270, loss: 0.0013310419162735343
step: 280, loss: 0.01590145379304886
step: 290, loss: 0.0010037326719611883
step: 300, loss: 0.0009036784758791327
step: 310, loss: 0.016201458871364594
step: 320, loss: 0.03856636583805084
step: 330, loss: 0.01841285638511181
step: 340, loss: 0.10369203984737396
step: 350, loss: 0.1268824189901352
step: 360, loss: 0.002648792928084731
step: 370, loss: 0.06749533116817474

```

```
step: 380, loss: 0.017005644738674164
step: 390, loss: 0.019129512831568718
step: 400, loss: 0.007285677827894688
step: 410, loss: 0.0005533366929739714
step: 420, loss: 0.051155950874090195
step: 430, loss: 0.00046102271880954504
step: 440, loss: 0.02103559859097004
step: 450, loss: 0.006410008296370506
step: 460, loss: 0.013336102478206158
step: 470, loss: 0.209971085190773
step: 480, loss: 0.005723627284169197
step: 490, loss: 0.0004078155034221709
step: 500, loss: 0.0015634649898856878
step: 510, loss: 0.00022137482301332057
step: 520, loss: 0.0066649471409618855
step: 530, loss: 0.006491576321423054
/usr/local/lib/python3.7/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=0.914180816888481, f1=0.9134001823154057,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00797764677554369
step: 10, loss: 0.029980454593896866
step: 20, loss: 0.005366603843867779
step: 30, loss: 0.0007417777669616044
step: 40, loss: 0.0038978306110948324
step: 50, loss: 0.004176388960331678
step: 60, loss: 0.14973534643650055
step: 70, loss: 0.01824156381189823
step: 80, loss: 0.011887572705745697
step: 90, loss: 0.039029691368341446
step: 100, loss: 0.04081528261303902
step: 110, loss: 0.01368462760001421
step: 120, loss: 0.001342979958280921
step: 130, loss: 0.020438319072127342
step: 140, loss: 0.015924308449029922
step: 150, loss: 0.000481471506645903
step: 160, loss: 0.0005098161636851728
step: 170, loss: 0.006021993700414896
step: 180, loss: 0.0030222805216908455
step: 190, loss: 0.018689556047320366
step: 200, loss: 0.003121537622064352
step: 210, loss: 0.02200811728835106
step: 220, loss: 0.06614386290311813
step: 230, loss: 0.0012222349178045988
```

```

step: 240, loss: 0.0041875699535012245
step: 250, loss: 0.06094181165099144
step: 260, loss: 0.0006769434548914433
step: 270, loss: 0.08880545943975449
step: 280, loss: 0.016484349966049194
step: 290, loss: 0.0020318084862083197
step: 300, loss: 0.0005845644045621157
step: 310, loss: 0.0006982056074775755
step: 320, loss: 0.01748116873204708
step: 330, loss: 0.000480919930851087
step: 340, loss: 0.00042406158172525465
step: 350, loss: 0.00010457188182044774
step: 360, loss: 0.0013641170226037502
step: 370, loss: 0.14958851039409637
step: 380, loss: 0.02340514212846756
step: 390, loss: 0.0029876481276005507
step: 400, loss: 0.0018921017181128263
step: 410, loss: 0.009874369949102402
step: 420, loss: 0.013337167911231518
step: 430, loss: 0.008839547634124756
step: 440, loss: 0.02096514217555523
step: 450, loss: 0.014235530979931355
step: 460, loss: 0.002115040086209774
step: 470, loss: 0.07337360829114914
step: 480, loss: 0.048775043338537216
step: 490, loss: 0.004322436172515154
step: 500, loss: 0.0013397142756730318
step: 510, loss: 0.0013633349444717169
step: 520, loss: 0.009984474629163742
step: 530, loss: 0.002104779938235879
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9166666666666666, f1=0.908584686774942,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002921480219811201
step: 10, loss: 0.0027494297828525305
step: 20, loss: 0.0019728231709450483
step: 30, loss: 0.011233135126531124
step: 40, loss: 0.0017822271911427379
step: 50, loss: 0.0006907671922817826
step: 60, loss: 0.0010910193668678403
step: 70, loss: 0.0025588874705135822
step: 80, loss: 0.0006112324772402644

```

```
step: 90, loss: 0.001057133194990456
step: 100, loss: 0.007777671795338392
step: 110, loss: 0.001905376324430108
step: 120, loss: 0.011881648562848568
step: 130, loss: 0.09909386187791824
step: 140, loss: 0.00014462452963925898
step: 150, loss: 0.002683323808014393
step: 160, loss: 0.012203048914670944
step: 170, loss: 0.08868023753166199
step: 180, loss: 0.0004896460450254381
step: 190, loss: 0.0024898704141378403
step: 200, loss: 0.0004452344437595457
step: 210, loss: 0.00201277039013803
step: 220, loss: 0.0004635325458366424
step: 230, loss: 0.001619155053049326
step: 240, loss: 0.006943858694285154
step: 250, loss: 0.024985091760754585
step: 260, loss: 0.11887705326080322
step: 270, loss: 0.0020380611531436443
step: 280, loss: 0.0002986258186865598
step: 290, loss: 0.005904079880565405
step: 300, loss: 0.00019822285685222596
step: 310, loss: 0.13123846054077148
step: 320, loss: 0.00013343036698643118
step: 330, loss: 0.0014043425908312201
step: 340, loss: 0.0019815792329609394
step: 350, loss: 0.00237883348017931
step: 360, loss: 0.02838197909295559
step: 370, loss: 0.0011166483163833618
step: 380, loss: 0.00780486548319459
step: 390, loss: 0.0012190036941319704
step: 400, loss: 0.04455104470252991
step: 410, loss: 0.001248067244887352
step: 420, loss: 7.411051046801731e-05
step: 430, loss: 0.00022183290275279433
step: 440, loss: 0.0021634879522025585
step: 450, loss: 0.0030763735994696617
step: 460, loss: 0.0016282653668895364
step: 470, loss: 0.0004902170039713383
step: 480, loss: 0.0025325785391032696
step: 490, loss: 0.00025156972697004676
step: 500, loss: 0.0011587167391553521
step: 510, loss: 0.00024559881421737373
step: 520, loss: 0.003455268917605281
step: 530, loss: 0.0010669715702533722
/usr/local/lib/python3.7/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=0.9142857142857143, f1=0.9163533834586467,  
best_f1=0.9225126088950022  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.0017139194533228874  
step: 10, loss: 0.051166992634534836  
step: 20, loss: 0.00018326500139664859  
step: 30, loss: 4.5223074266687036e-05  
step: 40, loss: 0.01923832669854164  
step: 50, loss: 0.0010307522024959326  
step: 60, loss: 0.00022264271683525294  
step: 70, loss: 0.04780334234237671  
step: 80, loss: 0.0020013339817523956  
step: 90, loss: 0.0030439174734055996  
step: 100, loss: 0.0005579938879236579  
step: 110, loss: 0.00258056353777647  
step: 120, loss: 0.0033132177777588367  
step: 130, loss: 0.00015439241542480886  
step: 140, loss: 0.008340012282133102  
step: 150, loss: 0.0015491110971197486  
step: 160, loss: 0.0005480190156958997  
step: 170, loss: 0.0002623538894113153  
step: 180, loss: 0.0001526332343928516  
step: 190, loss: 0.009466477669775486  
step: 200, loss: 0.0016241109697148204  
step: 210, loss: 0.005165624897927046  
step: 220, loss: 0.06130066141486168  
step: 230, loss: 0.0005486204172484577  
step: 240, loss: 0.03932417556643486  
step: 250, loss: 0.019165102392435074  
step: 260, loss: 0.0003729678282979876  
step: 270, loss: 0.01478502806276083  
step: 280, loss: 0.000257939042057842  
step: 290, loss: 0.00023774021246936172  
step: 300, loss: 0.0010624813148751855  
step: 310, loss: 0.01094732154160738  
step: 320, loss: 0.006340810563415289  
step: 330, loss: 0.0009770873002707958  
step: 340, loss: 0.0035828084219247103  
step: 350, loss: 0.00018942306633107364  
step: 360, loss: 0.03962403163313866  
step: 370, loss: 0.0033198888413608074  
step: 380, loss: 0.00020780217892024666  
step: 390, loss: 0.0010082381777465343  
step: 400, loss: 0.013597584329545498  
step: 410, loss: 0.00038319724262692034  
step: 420, loss: 0.0008233624394051731
```

```
step: 430, loss: 7.504112727474421e-05
step: 440, loss: 0.00010263719741487876
step: 450, loss: 0.00011451568570919335
step: 460, loss: 0.06269896030426025
step: 470, loss: 0.0001330096274614334
step: 480, loss: 0.000156574125867337
step: 490, loss: 0.00013467183453030884
step: 500, loss: 0.0034361029975116253
step: 510, loss: 0.0007939253700897098
step: 520, loss: 0.001136432751081884
step: 530, loss: 0.00025313906371593475
/usr/local/lib/python3.7/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=0.9160018770530267, f1=0.9128397375820055,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011987151810899377
step: 10, loss: 0.00016986859554890543
step: 20, loss: 0.0008991907816380262
step: 30, loss: 8.545596938347444e-05
step: 40, loss: 9.348469029646367e-05
step: 50, loss: 0.04216461256146431
step: 60, loss: 0.013005109503865242
step: 70, loss: 0.00044267348130233586
step: 80, loss: 0.0008801101357676089
step: 90, loss: 0.00041951125604100525
step: 100, loss: 0.0006084137712605298
step: 110, loss: 0.005032145418226719
step: 120, loss: 0.002364729531109333
step: 130, loss: 0.0003669265715871006
step: 140, loss: 0.0010210934560745955
step: 150, loss: 0.006372350733727217
step: 160, loss: 0.0013923079241067171
step: 170, loss: 0.0011329443659633398
step: 180, loss: 0.0005990497884340584
step: 190, loss: 0.0012848149053752422
step: 200, loss: 0.00019345773034729064
step: 210, loss: 0.0190170519053936
step: 220, loss: 0.00041590494220145047
step: 230, loss: 7.088868005666882e-05
step: 240, loss: 0.0001653251238167286
step: 250, loss: 0.00037165364483371377
step: 260, loss: 0.00030240530031733215
step: 270, loss: 0.0008662418113090098
step: 280, loss: 3.333587301312946e-05
```

```

step: 290, loss: 0.0005813025636598468
step: 300, loss: 0.0005272425478324294
step: 310, loss: 0.0002795149339362979
step: 320, loss: 0.0001953733153641224
step: 330, loss: 0.04534928500652313
step: 340, loss: 0.0392594113945961
step: 350, loss: 0.000307413749396801
step: 360, loss: 0.0012520031305029988
step: 370, loss: 0.0010145900305360556
step: 380, loss: 0.00022193822951521724
step: 390, loss: 6.238776404643431e-05
step: 400, loss: 0.000304053770378232
step: 410, loss: 0.0001268197229364887
step: 420, loss: 0.0027083877939730883
step: 430, loss: 0.00010507458500796929
step: 440, loss: 2.6735431674751453e-05
step: 450, loss: 7.58441528887488e-05
step: 460, loss: 7.661677955184132e-05
step: 470, loss: 9.988414967665449e-05
step: 480, loss: 0.00014282883785199374
step: 490, loss: 0.06875381618738174
step: 500, loss: 0.04642502963542938
step: 510, loss: 0.0040186503902077675
step: 520, loss: 0.006334066390991211
step: 530, loss: 0.00025606059352867305
/usr/local/lib/python3.7/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=0.9123951537744641, f1=0.9132462686567163,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000807680597063154
step: 10, loss: 0.008188356645405293
step: 20, loss: 0.0017416131449863315
step: 30, loss: 0.013174590654671192
step: 40, loss: 0.014123599976301193
step: 50, loss: 0.0011129004415124655
step: 60, loss: 0.00011627247295109555
step: 70, loss: 5.815941403852776e-05
step: 80, loss: 3.8106911233626306e-05
step: 90, loss: 0.003294155700132251
step: 100, loss: 3.609601117204875e-05
step: 110, loss: 0.0001628059835638851
step: 120, loss: 0.006662735715508461
step: 130, loss: 0.0013137846253812313
step: 140, loss: 9.293235780205578e-05

```



```

step: 150, loss: 8.054871432250366e-05
step: 160, loss: 0.039620839059352875
step: 170, loss: 0.044657353311777115
step: 180, loss: 6.0067115555284545e-05
step: 190, loss: 0.03325335681438446
step: 200, loss: 0.0010226726299151778
step: 210, loss: 5.532396608032286e-05
step: 220, loss: 0.0002848880540113896
step: 230, loss: 0.00042884040158241987
step: 240, loss: 3.91490466427058e-05
step: 250, loss: 0.004029050003737211
step: 260, loss: 3.680714871734381e-05
step: 270, loss: 0.011839240789413452
step: 280, loss: 0.002192357089370489
step: 290, loss: 0.001160724787041545
step: 300, loss: 0.007064009550958872
step: 310, loss: 0.014867139048874378
step: 320, loss: 0.013464173302054405
step: 330, loss: 0.0053657228127121925
step: 340, loss: 0.004023673012852669
step: 350, loss: 0.034084025770425797
step: 360, loss: 0.001998089486733079
step: 370, loss: 0.00032198824919760227
step: 380, loss: 0.00010140597441932186
step: 390, loss: 0.010206962004303932
step: 400, loss: 3.375994856469333e-05
step: 410, loss: 0.029663370922207832
step: 420, loss: 0.001564681064337492
step: 430, loss: 7.93458748375997e-05
step: 440, loss: 3.649159043561667e-05
step: 450, loss: 4.503508534980938e-05
step: 460, loss: 0.00875844806432724
step: 470, loss: 0.0013343188911676407
step: 480, loss: 0.008438076823949814
step: 490, loss: 0.001076755579560995
step: 500, loss: 0.00012417262769304216
step: 510, loss: 0.0016196678625419736
step: 520, loss: 6.512706750072539e-05
step: 530, loss: 5.353194137569517e-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=0.9125295508274233, f1=0.9147432878002827,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003943921532481909

```

step: 10, loss: 0.003697451204061508
step: 20, loss: 2.1647130779456347e-05
step: 30, loss: 3.494137854431756e-05
step: 40, loss: 0.00012841737770941108
step: 50, loss: 0.0015744883567094803
step: 60, loss: 8.172134403139353e-05
step: 70, loss: 0.0013434754218906164
step: 80, loss: 0.002486616838723421
step: 90, loss: 0.00014030389138497412
step: 100, loss: 0.00659843347966671
step: 110, loss: 0.0005226240027695894
step: 120, loss: 0.004027343355119228
step: 130, loss: 0.001798907178454101
step: 140, loss: 0.00044394363067112863
step: 150, loss: 0.0011300260666757822
step: 160, loss: 0.0007877878379076719
step: 170, loss: 0.0006980061298236251
step: 180, loss: 0.0015929306391626596
step: 190, loss: 9.490422235103324e-05
step: 200, loss: 0.00010175027273362502
step: 210, loss: 0.012689009308815002
step: 220, loss: 0.00011854183685500175
step: 230, loss: 0.00231803092174232
step: 240, loss: 0.0003175162710249424
step: 250, loss: 0.0001014387235045433
step: 260, loss: 0.00023888834402896464
step: 270, loss: 0.00458826869726181
step: 280, loss: 0.00013495999155566096
step: 290, loss: 0.0010552575113251805
step: 300, loss: 0.07707898318767548
step: 310, loss: 0.00048438552767038345
step: 320, loss: 0.07548798620700836
step: 330, loss: 0.00013445195509120822
step: 340, loss: 0.0011100209085270762
step: 350, loss: 0.01003547478467226
step: 360, loss: 0.00024078311980701983
step: 370, loss: 0.0005425909184850752
step: 380, loss: 0.00010729602217907086
step: 390, loss: 9.643468365538865e-05
step: 400, loss: 0.00020070104801561683
step: 410, loss: 0.0008277007145807147
step: 420, loss: 0.014116453006863594
step: 430, loss: 0.00016203071572817862
step: 440, loss: 0.0004746988997794688
step: 450, loss: 0.02946186251938343
step: 460, loss: 0.00129062426276505
step: 470, loss: 0.002435549395158887
step: 480, loss: 0.00018411035125609487

```
step: 490, loss: 0.00032229223870672286
step: 500, loss: 0.007027051877230406
step: 510, loss: 0.00956633035093546
step: 520, loss: 0.06091010570526123
step: 530, loss: 6.645637768087909e-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=0.9121184088806661, f1=0.9122645842903079,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023357797181233764
step: 10, loss: 5.138949563843198e-05
step: 20, loss: 0.007566753774881363
step: 30, loss: 0.009883658029139042
step: 40, loss: 0.0006473232642747462
step: 50, loss: 0.00013563230459112674
step: 60, loss: 6.13079100730829e-05
step: 70, loss: 0.0007718263659626245
step: 80, loss: 0.00039118726272135973
step: 90, loss: 0.0002719019539654255
step: 100, loss: 4.764162440551445e-05
step: 110, loss: 0.005674486979842186
step: 120, loss: 0.0002070326590910554
step: 130, loss: 0.0019809964578598738
step: 140, loss: 0.000619164842646569
step: 150, loss: 0.0012945131165906787
step: 160, loss: 0.0007774766418151557
step: 170, loss: 5.916965892538428e-05
step: 180, loss: 9.93323337752372e-05
step: 190, loss: 0.0003702342219185084
step: 200, loss: 0.0004390446993056685
step: 210, loss: 0.0003780547122005373
step: 220, loss: 0.00011095349327661097
step: 230, loss: 0.020802641287446022
step: 240, loss: 3.641329021775164e-05
step: 250, loss: 0.028818659484386444
step: 260, loss: 0.001364396302960813
step: 270, loss: 0.0005009659798815846
step: 280, loss: 0.00021133101836312562
step: 290, loss: 0.0006574599537998438
step: 300, loss: 0.00025724517763592303
step: 310, loss: 0.0005756831378675997
step: 320, loss: 5.570355642703362e-05
step: 330, loss: 0.008639597333967686
step: 340, loss: 6.833517545601353e-05
```

```
step: 350, loss: 0.0663762018084526
step: 360, loss: 6.947529618628323e-05
step: 370, loss: 0.004823804832994938
step: 380, loss: 0.0024051442742347717
step: 390, loss: 0.0005998257547616959
step: 400, loss: 0.00011280673061264679
step: 410, loss: 0.020717578008770943
step: 420, loss: 0.00016053739818744361
step: 430, loss: 8.141678699757904e-05
step: 440, loss: 0.04589226841926575
step: 450, loss: 0.0006549077806994319
step: 460, loss: 4.721659206552431e-05
step: 470, loss: 0.007724394556134939
step: 480, loss: 0.00018994195852428675
step: 490, loss: 0.0005753324367105961
step: 500, loss: 0.000883736414834857
step: 510, loss: 3.344801007187925e-05
step: 520, loss: 0.0001935924228746444
step: 530, loss: 6.659391510766e-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=0.9147727272727272, f1=0.9167058823529411,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015995465219020844
step: 10, loss: 0.0007162088295444846
step: 20, loss: 9.23764455365017e-05
step: 30, loss: 0.0012413420481607318
step: 40, loss: 0.0003888753126375377
step: 50, loss: 0.0008971237693913281
step: 60, loss: 0.0005131957004778087
step: 70, loss: 8.83020693436265e-05
step: 80, loss: 0.0003099565801676363
step: 90, loss: 7.35135399736464e-05
step: 100, loss: 5.923638309468515e-05
step: 110, loss: 9.520793537376449e-05
step: 120, loss: 0.0002175937406718731
step: 130, loss: 0.0009806769667193294
step: 140, loss: 0.009877721779048443
step: 150, loss: 0.006960081402212381
step: 160, loss: 0.0029357625171542168
step: 170, loss: 0.0005693384446203709
step: 180, loss: 0.0006964282365515828
step: 190, loss: 0.0003094766288995743
step: 200, loss: 0.00018115175771526992
```

```

step: 210, loss: 0.00024589220993220806
step: 220, loss: 0.0006176009774208069
step: 230, loss: 0.009739279747009277
step: 240, loss: 0.0004452868306543678
step: 250, loss: 0.000251550052780658
step: 260, loss: 4.731838998850435e-05
step: 270, loss: 0.007684269454330206
step: 280, loss: 7.335640111705288e-05
step: 290, loss: 0.0033326486591249704
step: 300, loss: 0.0027465764433145523
step: 310, loss: 0.00020381172362249345
step: 320, loss: 0.0009440452558919787
step: 330, loss: 0.005754212848842144
step: 340, loss: 9.365125151816756e-05
step: 350, loss: 0.002266878029331565
step: 360, loss: 0.0018604923970997334
step: 370, loss: 0.00017321757331956178
step: 380, loss: 0.0014999073464423418
step: 390, loss: 0.00015193820581771433
step: 400, loss: 2.4884471713448875e-05
step: 410, loss: 0.000153037064592354
step: 420, loss: 0.0026381469797343016
step: 430, loss: 0.0002787589037325233
step: 440, loss: 0.08142067492008209
step: 450, loss: 0.0002669976674951613
step: 460, loss: 0.0067437924444675446
step: 470, loss: 3.584712976589799e-05
step: 480, loss: 7.275442476384342e-05
step: 490, loss: 0.00010273020598106086
step: 500, loss: 0.00457654008641839
step: 510, loss: 0.0007237974205054343
step: 520, loss: 6.274793122429401e-05
step: 530, loss: 0.0002365978725720197
/usr/local/lib/python3.7/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=0.9161648911533117, f1=0.9178522257916476,
best_f1=0.9225126088950022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000301279011182487
step: 10, loss: 7.181651017162949e-05
step: 20, loss: 0.0009050974040292203
step: 30, loss: 0.0012435016687959433
step: 40, loss: 0.00012266392877791077
step: 50, loss: 0.00026709280791692436
step: 60, loss: 5.0541555538075045e-05

```

step: 70, loss: 0.0003205255197826773
step: 80, loss: 0.0001635228400118649
step: 90, loss: 7.165849820012227e-05
step: 100, loss: 8.50370415719226e-05
step: 110, loss: 0.004730027634650469
step: 120, loss: 0.0038393421564251184
step: 130, loss: 2.4210294213844463e-05
step: 140, loss: 0.0013251069467514753
step: 150, loss: 0.0011588167399168015
step: 160, loss: 0.00013457036402542144
step: 170, loss: 0.0007228457252494991
step: 180, loss: 0.00014478950470220298
step: 190, loss: 0.000678970362059772
step: 200, loss: 0.00010064840171253309
step: 210, loss: 0.0002173044194933027
step: 220, loss: 0.00010524818208068609
step: 230, loss: 0.0010661226697266102
step: 240, loss: 6.038865467417054e-05
step: 250, loss: 0.0006074074190109968
step: 260, loss: 0.0038903390523046255
step: 270, loss: 0.0010410358663648367
step: 280, loss: 0.00014635060506407171
step: 290, loss: 0.04041963815689087
step: 300, loss: 0.008861619979143143
step: 310, loss: 0.0008331384742632508
step: 320, loss: 0.01971755176782608
step: 330, loss: 0.0017712682019919157
step: 340, loss: 0.0009699140791781247
step: 350, loss: 0.0001344196789432317
step: 360, loss: 5.130098725203425e-05
step: 370, loss: 2.4117109205690213e-05
step: 380, loss: 0.0014771298738196492
step: 390, loss: 0.0002984637103509158
step: 400, loss: 0.00046860220027156174
step: 410, loss: 0.001080988091416657
step: 420, loss: 0.00023550847254227847
step: 430, loss: 6.192122964421287e-05
step: 440, loss: 0.0024532638490200043
step: 450, loss: 0.0007729424396529794
step: 460, loss: 3.98631309508346e-05
step: 470, loss: 4.1091698221862316e-05
step: 480, loss: 4.8105001042131335e-05
step: 490, loss: 0.0002556360268499702
step: 500, loss: 6.264111289056018e-05
step: 510, loss: 0.0001382610062137246
step: 520, loss: 0.000218727916944772
step: 530, loss: 3.827114778687246e-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=0.9167828916782892, f1=0.9184707508060802,
best_f1=0.9225126088950022
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.")
5742it [00:16, 348.94it/s]
load_f1 = 0.9230064161319891
real_f1 = 0.919634703196347
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 [00:11, 379.01it/s]

```