

ADirty_90_3_5_roberta

August 11, 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 | 8.54 MiB/s, done.
Resolving deltas: 100% (142/142), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.6 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 19.6 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 64.8 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.3 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 69.4 MB/s
Requirement already satisfied: sklearn==0.0 in
/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
```

```

Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
|                                     | 6.4 MB 23.6 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 | 17.62 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 26.0 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 55.4 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
|                                     | 753.4 MB 7.0 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
|                                     | 56 kB 5.5 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
|                                     | 884 kB 64.9 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 63.1 MB/s
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Collecting thinc<7.4.0,>=7.3.0
  Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
    | 2.2 MB 62.6 MB/s
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Collecting srsly<1.1.0,>=0.1.0
  Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
    | 184 kB 72.9 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
Requirement already satisfied: 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: 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: 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 68.5 MB/s
Collecting catalogue<1.1.0,>=0.0.7
  Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses

```

```

    Downloading sacremoses-0.0.53.tar.gz (880 kB)
      |                               | 880 kB 66.1 MB/s
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
      |                               | 3.0 MB 61.8 MB/s
Requirement already satisfied: filelock in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (3.7.1)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line 4)) (4.12.0)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (4.1.1)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-r requirements.txt (line
4)) (3.8.1)
Requirement already satisfied: 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=1449921
sha256=3b7e38f0fa692305fdd947b8f6016c9441f7d6f4ed2981d10d1f153dd343d1c5
    Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b28
1f323842bbda9bd819eb9a3ead
  Building wheel for sacremoses (setup.py) ... done
    Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl

```

```
size=895260
sha256=b4e6a17d3ac7d71baca0a9b56eda8d9574705a5aaab051829ec2fe0ffe589c73
  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: 10091, done.
remote: Counting objects: 100% (207/207), done.
remote: Compressing objects: 100% (144/144), done.
remote: Total 10091 (delta 100), reused 131 (delta 59), pack-reused 9884
Receiving objects: 100% (10091/10091), 14.96 MiB | 1.74 MiB/s, done.
Resolving deltas: 100% (6908/6908), 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-gvzxkpz3
Created temporary directory: /tmp/pip-req-tracker-9w6c3awy
Initialized build tracking at /tmp/pip-req-tracker-9w6c3awy
Created build tracker: /tmp/pip-req-tracker-9w6c3awy
Entered build tracker: /tmp/pip-req-tracker-9w6c3awy
```

```

Created temporary directory: /tmp/pip-install-durxbajx
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-0s163tny
  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-9w6c3awy'
  Running setup.py (path:/tmp/pip-req-build-0s163tny/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-tnb20hf6
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0


  running egg_info
  creating /tmp/pip-pip-egg-info-tnb20hf6/apex.egg-info
  writing /tmp/pip-pip-egg-info-tnb20hf6/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-tnb20hf6/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-tnb20hf6/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-tnb20hf6/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-tnb20hf6/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-0s163tny 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-9w6c3awy'
Created temporary directory: /tmp/pip-unpack-sotcm81e
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-a49jb_o1
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-a49jb_o1
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''''/tmp/pip-req-build-0s163tny/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-0s163tny/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-a49jb_o1
```

```
torch.__version__ = 1.4.0
```

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

```

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

```

```

build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/__init__.py ->
build/lib/apex/contrib/multihead_attn
  copying
apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
  copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
  creating build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
  copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
  creating build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
  copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
  creating build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
  copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
  creating build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/transducer.py ->
build/lib/apex/contrib/transducer
  copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
  creating build/lib/apex/contrib/peer_memory

```

```

    copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
    copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/__init__.py ->
build/lib/apex/contrib/conv_bias_relu
    copying apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/lib/apex/contrib/conv_bias_relu
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d

```

```

    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/layers
    copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
    copying apex/transformer/layers/layer_norm.py ->
build/lib/apex/transformer/layers
    creating build/lib/apex/transformer/functional
    copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
    copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
    creating build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/utils.py ->
build/lib/apex/transformer/pipeline_parallel
    creating build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->

```

```

build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/testing
  copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  creating build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  creating build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_with_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipeline_parallel_without_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipeline_parallel.py -> build/lib/apex/transformer/pipeline_parallel/schedules
  copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
  installing to build/bdist.linux-x86_64/wheel
  running install
  running install_lib
  creating build/bdist.linux-x86_64
  creating build/bdist.linux-x86_64/wheel
  creating build/bdist.linux-x86_64/wheel/apex
  creating build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_initialize.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/handle.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_amp_state.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/wrap.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/rnn_compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/_process_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
  copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/amp.py -> build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/frontend.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/compat.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  copying build/lib/apex/amp/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
  creating build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_func.py

```

```

-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
    creating build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_sgd.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
    creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib

```



```

    creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    creating build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_halo_exchange_module_tests.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    copying build/lib/apex/contrib/peer_memory/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
    creating build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    copying build/lib/apex/contrib/conv_bias_relu/conv_bias_relu.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/conv_bias_relu
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck

```

```

    copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
    creating build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    copying build/lib/apex/contrib/layer_norm/layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
    creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
    creating build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    creating build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->

```

```

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

```

```

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

```

```

    copying build/lib/apex/transformer/testing/standalone_gpt.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/arguments.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/log_util.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    running install_egg_info
    running egg_info
    creating apex.egg-info
    writing apex.egg-info/PKG-INFO
    writing dependency_links to apex.egg-info/dependency_links.txt
    writing top-level names to apex.egg-info/top_level.txt
    writing manifest file 'apex.egg-info/SOURCES.txt'
    adding license file 'LICENSE'
    writing manifest file 'apex.egg-info/SOURCES.txt'
    Copying apex.egg-info to build/bdist.linux-x86_64/wheel/apex-0.1-py3.7.egg-
info
    running install_scripts
    adding license file "LICENSE" (matched pattern "LICEN[CS]E*")
    creating build/bdist.linux-x86_64/wheel/apex-0.1.dist-info/WHEEL
    creating '/tmp/pip-wheel-a49jb_o1/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=b0c1c837dd6e515375ad25629070eaa487a999a993be2113a5671e921c429301
Stored in directory: /tmp/pip-ephem-wheel-cache-
gvzxkpz3/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-9w6c3awy'
/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 32.7 MB/s
Collecting awscli
  Downloading awscli-1.25.48-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 58.4 MB/s
Collecting botocore==1.27.48
  Downloading botocore-1.27.48-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 54.9 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 8.2 MB/s
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 63.2 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

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

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

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

Successfully installed awscli-1.25.48 botocore-1.27.48 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

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

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
    |                                     | 776.8 MB 17 kB/s
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 985, done.
remote: Counting objects: 100% (55/55), done.
remote: Compressing objects: 100% (51/51), done.
remote: Total 985 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (985/985), 252.10 MiB | 17.53 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1274/1274), 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/datasasErros/ADirty_90_3_5/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

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

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

Downloading: 100% 481/481 [00:00<00:00, 417kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 811kB/s]

Downloading: 100% 456k/456k [00:01<00:00, 416kB/s]

Downloading: 100% 501M/501M [00:07<00:00, 68.7MB/s]

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

Defaults for this optimization level are:

enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.41138964891433716
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=0.3076923076923077, f1=0.24444444444444444,
best_f1=0.24444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.43859201669692993
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3373493975903615, f1=0.25316455696202533,
best_f1=0.25316455696202533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4355158805847168
/usr/local/lib/python3.7/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=0.3076923076923077, f1=0.2823529411764706,
best_f1=0.25316455696202533
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24790231883525848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.37333333333333335, f1=0.30000000000000004,
best_f1=0.30000000000000004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2856398820877075
/usr/local/lib/python3.7/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=0.2916666666666667, f1=0.3010752688172043,
best_f1=0.30000000000000004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33419474959373474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3783783783783784, 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: 0.5935828685760498
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.3111111111111111, f1=0.2758620689655173,
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: 0.5075299143791199
/usr/local/lib/python3.7/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=0.3684210526315789, f1=0.36923076923076925,
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: 0.3876578211784363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.44444444444444445, f1=0.4090909090909091,
best_f1=0.4090909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.42664986848831177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4745762711864407, f1=0.42553191489361697,
best_f1=0.42553191489361697

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4746597707271576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.49122807017543857, f1=0.42553191489361697,
best_f1=0.42553191489361697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25290536880493164
/usr/local/lib/python3.7/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=0.45833333333333326, f1=0.5789473684210527,
best_f1=0.42553191489361697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2970203161239624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5185185185185185, f1=0.4761904761904762,
best_f1=0.4761904761904762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20660606026649475
/usr/local/lib/python3.7/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=0.5185185185185185, f1=0.43137254901960786,
best_f1=0.4761904761904762
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2977107763290405

```

```

/usr/local/lib/python3.7/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=0.5185185185185185, f1=0.43137254901960786,
best_f1=0.4761904761904762
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, 107606.90it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.5116279069767441
real_f1 = 0.5283018867924528
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:29, 147.19it/s]

```

1.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled           : True
```

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5594528317451477
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.426152765750885
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.4639163911342621
step: 30, loss: 0.28728801012039185
step: 40, loss: 0.32928267121315
step: 50, loss: 0.6241447329521179
step: 60, loss: 0.4424308240413666
step: 70, loss: 0.25717243552207947
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.502198338508606
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.3603366017341614
step: 100, loss: 0.10199589282274246
step: 110, loss: 0.2527065873146057
step: 120, loss: 0.19465531408786774
step: 130, loss: 0.016640830785036087
step: 140, loss: 0.07389167696237564
step: 150, loss: 0.1039348915219307
step: 160, loss: 0.051416635513305664
step: 170, loss: 0.22984512150287628
step: 180, loss: 0.02081911824643612
step: 190, loss: 0.13898733258247375
step: 200, loss: 0.07074031978845596
step: 210, loss: 0.01985827088356018
step: 220, loss: 0.037789300084114075
step: 230, loss: 0.03189815580844879
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9751131221719457, f1=0.9680365296803655,

```

```

best_f1=0.9680365296803655
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0029032211750745773
step: 10, loss: 0.10406063497066498
step: 20, loss: 0.020126042887568474
step: 30, loss: 0.01611391268670559
step: 40, loss: 0.015947800129652023
step: 50, loss: 0.004891406279057264
step: 60, loss: 0.0067612421698868275
step: 70, loss: 0.021271973848342896
step: 80, loss: 0.08145947009325027
step: 90, loss: 0.010733840987086296
step: 100, loss: 0.007137634791433811
step: 110, loss: 0.17646963894367218
step: 120, loss: 0.0016096564941108227
step: 130, loss: 0.0027878712862730026
step: 140, loss: 0.0076788137666881084
step: 150, loss: 0.11563393473625183
step: 160, loss: 0.007876909337937832
step: 170, loss: 0.08905985951423645
step: 180, loss: 0.008960344828665257
step: 190, loss: 0.11518072336912155
step: 200, loss: 0.009689787402749062
step: 210, loss: 0.11641671508550644
step: 220, loss: 0.02188892476260662
step: 230, loss: 0.14674246311187744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9773755656108598,
best_f1=0.9773755656108598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005520571488887072
step: 10, loss: 0.033481717109680176
step: 20, loss: 0.10210629552602768
step: 30, loss: 0.0019659230019897223
step: 40, loss: 0.008607838302850723
step: 50, loss: 0.04264243692159653
step: 60, loss: 0.011702976189553738
step: 70, loss: 0.0019984757527709007

```

```

step: 80, loss: 0.0008596020052209496
step: 90, loss: 0.005943822208791971
step: 100, loss: 0.0027064955793321133
step: 110, loss: 0.0035992946941405535
step: 120, loss: 0.0006418082630261779
step: 130, loss: 0.002146809594705701
step: 140, loss: 0.00046450516674667597
step: 150, loss: 0.057408738881349564
step: 160, loss: 0.0020453904289752245
step: 170, loss: 0.0018004290759563446
step: 180, loss: 0.08300425857305527
step: 190, loss: 0.010702645406126976
step: 200, loss: 0.00787993986159563
step: 210, loss: 0.0021690744906663895
step: 220, loss: 0.006581088528037071
step: 230, loss: 0.025018360465765
/usr/local/lib/python3.7/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=0.9783845278725825, f1=0.9759999999999999,
best_f1=0.9773755656108598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012682436965405941
step: 10, loss: 0.0011268671369180083
step: 20, loss: 0.0012164213694632053
step: 30, loss: 0.003400753950700164
step: 40, loss: 0.02341885305941105
step: 50, loss: 0.005095299798995256
step: 60, loss: 0.0011905863648280501
step: 70, loss: 0.0015707946149632335
step: 80, loss: 0.0013853221898898482
step: 90, loss: 0.002246077638119459
step: 100, loss: 0.0073193134739995
step: 110, loss: 0.002807039301842451
step: 120, loss: 0.0017775368178263307
step: 130, loss: 0.011494860053062439
step: 140, loss: 0.0003693613689392805
step: 150, loss: 0.0003195127646904439
step: 160, loss: 0.0005670480313710868
step: 170, loss: 0.005471720825880766
step: 180, loss: 0.09682486206293106
step: 190, loss: 0.004555656109005213
step: 200, loss: 0.017059743404388428
step: 210, loss: 0.0009029322536662221
step: 220, loss: 0.049264129251241684
step: 230, loss: 0.007244246546179056

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854096520763187, f1=0.9820224719101124,
best_f1=0.9820224719101124
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02619859203696251
step: 10, loss: 0.002304235240444541
step: 20, loss: 0.001323841162957251
step: 30, loss: 0.0015646599931642413
step: 40, loss: 0.0008431965834461153
step: 50, loss: 0.0006622375803999603
step: 60, loss: 0.0008432514150626957
step: 70, loss: 0.001037242473103106
step: 80, loss: 0.03463221713900566
step: 90, loss: 0.016542337834835052
step: 100, loss: 0.00022278103278949857
step: 110, loss: 0.001418804400600493
step: 120, loss: 0.21143901348114014
step: 130, loss: 0.0005831983289681375
step: 140, loss: 0.00105130928568542
step: 150, loss: 0.06033499166369438
step: 160, loss: 0.0009175627492368221
step: 170, loss: 0.0020632471423596144
step: 180, loss: 0.02343612350523472
step: 190, loss: 0.006931660696864128
step: 200, loss: 0.03379898890852928
step: 210, loss: 0.021391648799180984
step: 220, loss: 0.0027216284070163965
step: 230, loss: 0.005947648547589779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.990990990990991, f1=0.9909706546275394,
best_f1=0.9909706546275394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0012401659041643143
step: 10, loss: 0.0041468762792646885
step: 20, loss: 0.002925332635641098
step: 30, loss: 0.0045775240287184715
step: 40, loss: 0.0006006951443850994
step: 50, loss: 0.0005530991475097835
step: 60, loss: 0.21601849794387817
step: 70, loss: 0.0017968651372939348
step: 80, loss: 0.054895661771297455
step: 90, loss: 0.010730482637882233
step: 100, loss: 0.001978188520297408
step: 110, loss: 0.000945572683122009
step: 120, loss: 0.0003612501604948193
step: 130, loss: 0.0006637598853558302
step: 140, loss: 0.0014650062657892704
step: 150, loss: 0.0005185747286304832
step: 160, loss: 0.043115030974149704
step: 170, loss: 0.004472588654607534
step: 180, loss: 0.00017103897698689252
step: 190, loss: 0.0005409217556007206
step: 200, loss: 0.014420310966670513
step: 210, loss: 0.00147503730840981
step: 220, loss: 0.015035044401884079
step: 230, loss: 0.0007823532214388251
/usr/local/lib/python3.7/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=0.9876819708846584, f1=0.9797752808988766,
best_f1=0.9909706546275394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009342344710603356
step: 10, loss: 0.0005284327198751271
step: 20, loss: 0.0020071968901902437
step: 30, loss: 0.00020591144857462496
step: 40, loss: 0.0003810326161328703
step: 50, loss: 0.0004949304857291281
step: 60, loss: 0.0006708185537718236
step: 70, loss: 0.00028071997803635895
step: 80, loss: 0.0012372353812679648
step: 90, loss: 0.0008579103159718215
step: 100, loss: 0.00039129555807448924
step: 110, loss: 0.0016003830824047327
step: 120, loss: 0.0008676068973727524
step: 130, loss: 0.0004212207277305424
step: 140, loss: 0.00011508754687383771
step: 150, loss: 0.0003610964340623468

```



```

step: 160, loss: 0.00022388403885997832
step: 170, loss: 0.00029179328703321517
step: 180, loss: 0.03461304306983948
step: 190, loss: 0.00022872694535180926
step: 200, loss: 0.004119702614843845
step: 210, loss: 0.007798140402883291
step: 220, loss: 0.0004366521316114813
step: 230, loss: 0.001008384395390749
/usr/local/lib/python3.7/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=0.9876543209876544, f1=0.9831649831649831,
best_f1=0.9909706546275394
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007590728346258402
step: 10, loss: 0.00671972893178463
step: 20, loss: 0.007098160218447447
step: 30, loss: 0.002171957166865468
step: 40, loss: 0.0013649447355419397
step: 50, loss: 0.001670626807026565
step: 60, loss: 0.004750428255647421
step: 70, loss: 0.00020986419986002147
step: 80, loss: 0.001622675103135407
step: 90, loss: 0.00017865611880552024
step: 100, loss: 0.00021758349612355232
step: 110, loss: 0.0006946335779502988
step: 120, loss: 0.00010319525608792901
step: 130, loss: 0.0002971123030874878
step: 140, loss: 0.00026923793484456837
step: 150, loss: 0.008807668462395668
step: 160, loss: 0.000633751624263823
step: 170, loss: 0.08422993123531342
step: 180, loss: 0.0008898893720470369
step: 190, loss: 0.0004692889051511884
step: 200, loss: 0.00031224131816998124
step: 210, loss: 0.0012246964033693075
step: 220, loss: 0.0029312586411833763
step: 230, loss: 0.0025388621725142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910112359550561, f1=0.9831649831649831,

```

```

best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007624350255355239
step: 10, loss: 0.0023321642074733973
step: 20, loss: 0.015668055042624474
step: 30, loss: 0.0004852791898883879
step: 40, loss: 0.00022163828543853015
step: 50, loss: 0.00033880164846777916
step: 60, loss: 0.00017195298278238624
step: 70, loss: 0.033761437982320786
step: 80, loss: 0.0015297499485313892
step: 90, loss: 0.01397309172898531
step: 100, loss: 0.002087402855977416
step: 110, loss: 0.0003333617059979588
step: 120, loss: 0.0002675863215699792
step: 130, loss: 0.0008090949268080294
step: 140, loss: 0.003252160269767046
step: 150, loss: 0.0011368909617885947
step: 160, loss: 0.08157844096422195
step: 170, loss: 0.0018952219979837537
step: 180, loss: 0.010100231505930424
step: 190, loss: 0.0003028852224815637
step: 200, loss: 0.0005937201785854995
step: 210, loss: 0.0028841898310929537
step: 220, loss: 0.0002490011975169182
step: 230, loss: 7.384144555544481e-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=0.9863636363636363, f1=0.9829738933030647,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020241799938958138
step: 10, loss: 0.0002114108792738989
step: 20, loss: 0.00039399502566084266
step: 30, loss: 0.00045309419510886073
step: 40, loss: 0.0070190271362662315
step: 50, loss: 0.0004046616959385574
step: 60, loss: 0.0008885483839549124
step: 70, loss: 0.0006034501711837947
step: 80, loss: 0.0001969316363101825
step: 90, loss: 0.00026698861620388925
step: 100, loss: 0.0002005032147280872
step: 110, loss: 0.001284239930100739

```

```

step: 120, loss: 0.00019558638450689614
step: 130, loss: 0.00029641471337527037
step: 140, loss: 0.0001369398960378021
step: 150, loss: 0.00036872801138088107
step: 160, loss: 0.00010732188820838928
step: 170, loss: 0.00017600294086150825
step: 180, loss: 0.004459569230675697
step: 190, loss: 0.0001950790756382048
step: 200, loss: 0.000518154411111027
step: 210, loss: 0.002999034011736512
step: 220, loss: 0.003229474416002631
step: 230, loss: 7.494715828215703e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.9910313901345291, f1=0.9865168539325843,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.0892507715616375e-05
step: 10, loss: 0.00012038202839903533
step: 20, loss: 9.653306915424764e-05
step: 30, loss: 7.404074858641252e-05
step: 40, loss: 3.07720220007468e-05
step: 50, loss: 4.5689361286349595e-05
step: 60, loss: 0.09002037346363068
step: 70, loss: 0.018080605193972588
step: 80, loss: 0.0003767946909647435
step: 90, loss: 0.0002918145328294486
step: 100, loss: 0.00018504432227928191
step: 110, loss: 0.0008259626338258386
step: 120, loss: 9.013589442474768e-05
step: 130, loss: 8.87203641468659e-05
step: 140, loss: 0.0003739367239177227
step: 150, loss: 0.0018496226985007524
step: 160, loss: 0.007907778955996037
step: 170, loss: 0.000190995866432786
step: 180, loss: 0.00012158403842477128
step: 190, loss: 0.010987343266606331
step: 200, loss: 0.00032150078914128244
step: 210, loss: 0.00019303348381072283
step: 220, loss: 0.00113511027302593
step: 230, loss: 0.00010570452286629006

```

```

/usr/local/lib/python3.7/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=0.987709497206704, f1=0.9821029082774049,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010557177301961929
step: 10, loss: 0.00010122002277057618
step: 20, loss: 0.005081820301711559
step: 30, loss: 0.031203050166368484
step: 40, loss: 0.0009209828567691147
step: 50, loss: 0.000844803114887327
step: 60, loss: 0.0007604074780829251
step: 70, loss: 0.00024774103076197207
step: 80, loss: 0.00012747812434099615
step: 90, loss: 0.0007871088455431163
step: 100, loss: 0.00021726716659031808
step: 110, loss: 0.00013501386274583638
step: 120, loss: 0.0001232087379321456
step: 130, loss: 8.67171329446137e-05
step: 140, loss: 0.00011169586650794372
step: 150, loss: 9.115792636293918e-05
step: 160, loss: 0.000600694096647203
step: 170, loss: 0.000703939818777144
step: 180, loss: 9.087631769943982e-05
step: 190, loss: 0.0005576623370870948
step: 200, loss: 0.00017533850041218102
step: 210, loss: 0.00014008361904416233
step: 220, loss: 5.644539851346053e-05
step: 230, loss: 0.00035540293902158737
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9853768278965129, f1=0.984304932735426,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.74084697291255e-05
step: 10, loss: 0.00012886915646959096
step: 20, loss: 9.616361057851464e-05
step: 30, loss: 0.0014257762813940644
step: 40, loss: 0.00018049831851385534
step: 50, loss: 0.0003888139617629349
step: 60, loss: 0.0002726053062360734
step: 70, loss: 0.00017314814613200724

```

```

step: 80, loss: 0.00028615043265745044
step: 90, loss: 7.797110447427258e-05
step: 100, loss: 0.0002917555102612823
step: 110, loss: 0.0002410035376669839
step: 120, loss: 0.0001687566691543907
step: 130, loss: 0.00022166708367876709
step: 140, loss: 8.098284888546914e-05
step: 150, loss: 4.910533607471734e-05
step: 160, loss: 0.0018039002316072583
step: 170, loss: 0.0001656301465118304
step: 180, loss: 0.0018093022517859936
step: 190, loss: 0.00010552906314842403
step: 200, loss: 6.584631773876026e-05
step: 210, loss: 0.00014314329018816352
step: 220, loss: 0.00021916895639151335
step: 230, loss: 6.30016002105549e-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=0.9876819708846584, f1=0.9832026875699889,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.347344606183469e-05
step: 10, loss: 7.933519373182207e-05
step: 20, loss: 5.3676350944442675e-05
step: 30, loss: 0.00012113719276385382
step: 40, loss: 0.00016146511188708246
step: 50, loss: 4.549498771666549e-05
step: 60, loss: 8.729195542400703e-05
step: 70, loss: 0.00019652668561320752
step: 80, loss: 5.274168142932467e-05
step: 90, loss: 0.00014842583914287388
step: 100, loss: 0.00010168866720050573
step: 110, loss: 7.135736814234406e-05
step: 120, loss: 1.809648711059708e-05
step: 130, loss: 0.00014514957729261369
step: 140, loss: 8.108558540698141e-05
step: 150, loss: 6.761275290045887e-05
step: 160, loss: 0.0011290640104562044
step: 170, loss: 9.14677293621935e-05
step: 180, loss: 0.00015548712690360844
step: 190, loss: 0.00013112569286022335
step: 200, loss: 0.00048414795310236514
step: 210, loss: 4.7458110202569515e-05
step: 220, loss: 0.0001286850165342912
step: 230, loss: 5.460292231873609e-05

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9887640449438202, f1=0.9854096520763187,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020373391453176737
step: 10, loss: 5.702055568690412e-05
step: 20, loss: 0.0008444447303190827
step: 30, loss: 6.0059435782022774e-05
step: 40, loss: 4.1487321141175926e-05
step: 50, loss: 5.056126246927306e-05
step: 60, loss: 0.04105804115533829
step: 70, loss: 5.9942263760603964e-05
step: 80, loss: 5.89656010561157e-05
step: 90, loss: 7.35834619263187e-05
step: 100, loss: 3.483234468149021e-05
step: 110, loss: 0.00022670128964819014
step: 120, loss: 0.0002774225431494415
step: 130, loss: 8.948156755650416e-05
step: 140, loss: 0.004575501196086407
step: 150, loss: 0.00044725966290570796
step: 160, loss: 0.00037151912692934275
step: 170, loss: 2.8779262720490806e-05
step: 180, loss: 0.00011273184645688161
step: 190, loss: 4.94705636810977e-05
step: 200, loss: 0.00015538488514721394
step: 210, loss: 0.00013019415200687945
step: 220, loss: 0.00014397522318176925
step: 230, loss: 9.1987443738617e-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=0.9887640449438202, f1=0.9865168539325843,
best_f1=0.9865168539325843
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:16, 149.88it/s]
load_f1 = 0.9898534385569334
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 [00:30, 143.66it/s]

```

1.1.3 DBLP-GoogleScholar - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

step: 0, loss: 0.6351011395454407

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

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

step: 10, loss: 0.38612601161003113

step: 20, loss: 0.3133067190647125

step: 30, loss: 0.36976903676986694

step: 40, loss: 0.37451326847076416

step: 50, loss: 0.6486261487007141
step: 60, loss: 0.33434930443763733
step: 70, loss: 0.4712095856666565
step: 80, loss: 0.4840647578239441
step: 90, loss: 0.42971599102020264
step: 100, loss: 0.5336258411407471
step: 110, loss: 0.48330163955688477
step: 120, loss: 0.5849539041519165
step: 130, loss: 0.5600461363792419
step: 140, loss: 0.369144082069397
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 150, loss: 0.40193066000938416
step: 160, loss: 0.49846264719963074
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 170, loss: 0.19392763078212738
step: 180, loss: 0.20836257934570312
step: 190, loss: 0.25089001655578613
step: 200, loss: 0.5663240551948547
step: 210, loss: 0.15442998707294464
step: 220, loss: 0.1682528853416443
step: 230, loss: 0.36979812383651733
step: 240, loss: 0.08767421543598175
step: 250, loss: 0.08031388372182846
step: 260, loss: 0.47808197140693665
step: 270, loss: 0.2268693596124649
step: 280, loss: 0.11248474568128586
step: 290, loss: 0.14444176852703094
step: 300, loss: 0.19852089881896973
step: 310, loss: 0.3766535222530365
step: 320, loss: 0.13476771116256714
step: 330, loss: 0.1244756430387497
step: 340, loss: 0.49138346314430237
step: 350, loss: 0.2982802391052246
step: 360, loss: 0.0627615749835968
step: 370, loss: 0.022944042459130287
step: 380, loss: 0.14215002954006195
step: 390, loss: 0.03650803491473198
step: 400, loss: 0.01734234020113945
step: 410, loss: 0.2714202404022217
step: 420, loss: 0.10159015655517578
step: 430, loss: 0.10040714591741562
step: 440, loss: 0.17616361379623413
step: 450, loss: 0.10390278697013855
step: 460, loss: 0.10853704810142517
step: 470, loss: 0.09111225605010986
step: 480, loss: 0.1851404756307602
step: 490, loss: 0.1374576985836029
step: 500, loss: 0.13309070467948914

```

step: 510, loss: 0.11793189495801926
step: 520, loss: 0.7279350161552429
step: 530, loss: 0.10063080489635468
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8838916934373566, f1=0.8821362799263353,
best_f1=0.8821362799263353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14847244322299957
step: 10, loss: 0.12544135749340057
step: 20, loss: 0.19450081884860992
step: 30, loss: 0.1987629383802414
step: 40, loss: 0.10997682064771652
step: 50, loss: 0.18963319063186646
step: 60, loss: 0.09831631183624268
step: 70, loss: 0.11352141201496124
step: 80, loss: 0.11711686849594116
step: 90, loss: 0.0455399714410305
step: 100, loss: 0.18694248795509338
step: 110, loss: 0.1352463960647583
step: 120, loss: 0.16131041944026947
step: 130, loss: 0.03046775422990322
step: 140, loss: 0.1611226350069046
step: 150, loss: 0.08304721117019653
step: 160, loss: 0.031019030138850212
step: 170, loss: 0.11304184049367905
step: 180, loss: 0.06660997122526169
step: 190, loss: 0.057459309697151184
step: 200, loss: 0.24896420538425446
step: 210, loss: 0.20394110679626465
step: 220, loss: 0.009946306236088276
step: 230, loss: 0.10693851113319397
step: 240, loss: 0.06275536119937897
step: 250, loss: 0.15281665325164795
step: 260, loss: 0.2625812888145447
step: 270, loss: 0.05904730409383774
step: 280, loss: 0.14349530637264252
step: 290, loss: 0.08465474098920822
step: 300, loss: 0.18484243750572205
step: 310, loss: 0.10333912819623947
step: 320, loss: 0.1947673261165619

```

```

step: 330, loss: 0.11911077797412872
step: 340, loss: 0.2228899449110031
step: 350, loss: 0.02836088091135025
step: 360, loss: 0.08518368750810623
step: 370, loss: 0.013317195698618889
step: 380, loss: 0.1502322405576706
step: 390, loss: 0.0706862360239029
step: 400, loss: 0.07863053679466248
step: 410, loss: 0.07792431116104126
step: 420, loss: 0.04500935971736908
step: 430, loss: 0.3473028838634491
step: 440, loss: 0.048050351440906525
step: 450, loss: 0.0707370713353157
step: 460, loss: 0.06472722440958023
step: 470, loss: 0.17852507531642914
step: 480, loss: 0.09556399285793304
step: 490, loss: 0.0717795342206955
step: 500, loss: 0.016396308317780495
step: 510, loss: 0.1361425518989563
step: 520, loss: 0.4722483456134796
step: 530, loss: 0.11091696470975876
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8866171003717472, f1=0.8973418881759854,
best_f1=0.8973418881759854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22811955213546753
step: 10, loss: 0.21165207028388977
step: 20, loss: 0.03481403365731239
step: 30, loss: 0.14381353557109833
step: 40, loss: 0.25117161870002747
step: 50, loss: 0.07211744040250778
step: 60, loss: 0.06654778122901917
step: 70, loss: 0.127228781580925
step: 80, loss: 0.06278495490550995
step: 90, loss: 0.04213269054889679
step: 100, loss: 0.09881757944822311
step: 110, loss: 0.09184791147708893
step: 120, loss: 0.20226092636585236
step: 130, loss: 0.1100887656211853
step: 140, loss: 0.0474117211997509

```

```

step: 150, loss: 0.08832366019487381
step: 160, loss: 0.08886843919754028
step: 170, loss: 0.012309521436691284
step: 180, loss: 0.06767550110816956
step: 190, loss: 0.12314288318157196
step: 200, loss: 0.029904361814260483
step: 210, loss: 0.18387414515018463
step: 220, loss: 0.04390004277229309
step: 230, loss: 0.058404017239809036
step: 240, loss: 0.11391881853342056
step: 250, loss: 0.20566126704216003
step: 260, loss: 0.026706982403993607
step: 270, loss: 0.05882973596453667
step: 280, loss: 0.010946234688162804
step: 290, loss: 0.05028095841407776
step: 300, loss: 0.10829918086528778
step: 310, loss: 0.23936814069747925
step: 320, loss: 0.08068688213825226
step: 330, loss: 0.042075015604496
step: 340, loss: 0.0174017995595932
step: 350, loss: 0.1449311226606369
step: 360, loss: 0.05544956773519516
step: 370, loss: 0.11404254287481308
step: 380, loss: 0.07844426482915878
step: 390, loss: 0.037922345101833344
step: 400, loss: 0.1806066781282425
step: 410, loss: 0.02471070922911167
step: 420, loss: 0.17063257098197937
step: 430, loss: 0.07564198970794678
step: 440, loss: 0.27632656693458557
step: 450, loss: 0.1497832089662552
step: 460, loss: 0.11089959740638733
step: 470, loss: 0.036705195903778076
step: 480, loss: 0.13683074712753296
step: 490, loss: 0.03289215639233589
step: 500, loss: 0.04977305233478546
step: 510, loss: 0.018474865704774857
step: 520, loss: 0.008796379901468754
step: 530, loss: 0.07201908528804779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9093432007400555, f1=0.9148642429820525,
best_f1=0.9148642429820525

```

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

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

```
step: 0, loss: 0.05245935916900635  
step: 10, loss: 0.02824772708117962  
step: 20, loss: 0.13919541239738464  
step: 30, loss: 0.0591355599462986  
step: 40, loss: 0.07408786565065384  
step: 50, loss: 0.06578163802623749  
step: 60, loss: 0.033786043524742126  
step: 70, loss: 0.10769366472959518  
step: 80, loss: 0.15578103065490723  
step: 90, loss: 0.1830172836780548  
step: 100, loss: 0.010826620273292065  
step: 110, loss: 0.12243637442588806  
step: 120, loss: 0.06441392004489899  
step: 130, loss: 0.10451457649469376  
step: 140, loss: 0.12075409293174744  
step: 150, loss: 0.037376370280981064  
step: 160, loss: 0.05559299886226654  
step: 170, loss: 0.10726892948150635  
step: 180, loss: 0.13734321296215057  
step: 190, loss: 0.2520022988319397  
step: 200, loss: 0.03347938880324364  
step: 210, loss: 0.03522009029984474  
step: 220, loss: 0.059637442231178284  
step: 230, loss: 0.012157713063061237  
step: 240, loss: 0.031465355306863785  
step: 250, loss: 0.1268928498029709  
step: 260, loss: 0.10228607058525085  
step: 270, loss: 0.18526417016983032  
step: 280, loss: 0.05470636114478111  
step: 290, loss: 0.05034111812710762  
step: 300, loss: 0.04994826018810272  
step: 310, loss: 0.012103640474379063  
step: 320, loss: 0.17352239787578583  
step: 330, loss: 0.037681519985198975  
step: 340, loss: 0.005946212448179722  
step: 350, loss: 0.21464702486991882  
step: 360, loss: 0.1007341593503952  
step: 370, loss: 0.012872946448624134  
step: 380, loss: 0.011543791741132736  
step: 390, loss: 0.011076614260673523  
step: 400, loss: 0.0162592064589262  
step: 410, loss: 0.007909908890724182  
step: 420, loss: 0.08768895268440247  
step: 430, loss: 0.023157037794589996  
step: 440, loss: 0.028019804507493973
```

```

step: 450, loss: 0.05359445512294769
step: 460, loss: 0.15735699236392975
step: 470, loss: 0.03426113724708557
step: 480, loss: 0.020692570134997368
step: 490, loss: 0.04787531495094299
step: 500, loss: 0.12325131893157959
step: 510, loss: 0.13307785987854004
step: 520, loss: 0.05611758306622505
step: 530, loss: 0.14107175171375275
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9178338001867413, f1=0.9182624941616068,
best_f1=0.9182624941616068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007968500256538391
step: 10, loss: 0.01027492806315422
step: 20, loss: 0.019204219803214073
step: 30, loss: 0.133038729429245
step: 40, loss: 0.06580730527639389
step: 50, loss: 0.10366249084472656
step: 60, loss: 0.017007512971758842
step: 70, loss: 0.016540059819817543
step: 80, loss: 0.008497867733240128
step: 90, loss: 0.013056902214884758
step: 100, loss: 0.11167081445455551
step: 110, loss: 0.09875364601612091
step: 120, loss: 0.12625162303447723
step: 130, loss: 0.020596042275428772
step: 140, loss: 0.01999702677130699
step: 150, loss: 0.02961566485464573
step: 160, loss: 0.08131398260593414
step: 170, loss: 0.04857851192355156
step: 180, loss: 0.0058292970061302185
step: 190, loss: 0.012386539950966835
step: 200, loss: 0.006401561200618744
step: 210, loss: 0.14196732640266418
step: 220, loss: 0.033282212913036346
step: 230, loss: 0.04085564985871315
step: 240, loss: 0.015318794175982475
step: 250, loss: 0.08338861912488937
step: 260, loss: 0.0020017530769109726

```

```

step: 270, loss: 0.009806566871702671
step: 280, loss: 0.04637272655963898
step: 290, loss: 0.02051803283393383
step: 300, loss: 0.0676451250910759
step: 310, loss: 0.11709047853946686
step: 320, loss: 0.17348837852478027
step: 330, loss: 0.026497818529605865
step: 340, loss: 0.03998910263180733
step: 350, loss: 0.016538865864276886
step: 360, loss: 0.0012428724439814687
step: 370, loss: 0.01726483181118965
step: 380, loss: 0.012271610088646412
step: 390, loss: 0.2376711666584015
step: 400, loss: 0.038892727345228195
step: 410, loss: 0.16916389763355255
step: 420, loss: 0.14851273596286774
step: 430, loss: 0.1474739909172058
step: 440, loss: 0.014888681471347809
step: 450, loss: 0.01605718955397606
step: 460, loss: 0.05502030625939369
step: 470, loss: 0.0455644316971302
step: 480, loss: 0.08203010261058807
step: 490, loss: 0.09451986104249954
step: 500, loss: 0.01985687017440796
step: 510, loss: 0.013070452027022839
step: 520, loss: 0.23676557838916779
step: 530, loss: 0.024433841928839684
/usr/local/lib/python3.7/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=0.9132893496701224, f1=0.912, best_f1=0.9182624941616068
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04616393893957138
step: 10, loss: 0.005283580161631107
step: 20, loss: 0.014713317155838013
step: 30, loss: 0.005029469262808561
step: 40, loss: 0.0371866337954998
step: 50, loss: 0.00238743401132524
step: 60, loss: 0.006975526921451092
step: 70, loss: 0.0016796945128589869
step: 80, loss: 0.01799045130610466
step: 90, loss: 0.03376137465238571
step: 100, loss: 0.28636252880096436
step: 110, loss: 0.02454954758286476
step: 120, loss: 0.05036143213510513
step: 130, loss: 0.005504448898136616

```

```

step: 140, loss: 0.005276423413306475
step: 150, loss: 0.006164176855236292
step: 160, loss: 0.028516534715890884
step: 170, loss: 0.003669972298666835
step: 180, loss: 0.03503444790840149
step: 190, loss: 0.29769033193588257
step: 200, loss: 0.05725831538438797
step: 210, loss: 0.0416690967977047
step: 220, loss: 0.009449665434658527
step: 230, loss: 0.007787223905324936
step: 240, loss: 0.045777201652526855
step: 250, loss: 0.19583937525749207
step: 260, loss: 0.009511456824839115
step: 270, loss: 0.016454022377729416
step: 280, loss: 0.011760985478758812
step: 290, loss: 0.0036420042160898447
step: 300, loss: 0.0052647097036242485
step: 310, loss: 0.046893078833818436
step: 320, loss: 0.005053083412349224
step: 330, loss: 0.009030308574438095
step: 340, loss: 0.059837453067302704
step: 350, loss: 0.025057530030608177
step: 360, loss: 0.06780587136745453
step: 370, loss: 0.004759653005748987
step: 380, loss: 0.006626870483160019
step: 390, loss: 0.004992565605789423
step: 400, loss: 0.035105518996715546
step: 410, loss: 0.0027090057265013456
step: 420, loss: 0.0009753339691087604
step: 430, loss: 0.017801212146878242
step: 440, loss: 0.0009980102768167853
step: 450, loss: 0.1379779428243637
step: 460, loss: 0.002115562791004777
step: 470, loss: 0.006441256031394005
step: 480, loss: 0.007757776882499456
step: 490, loss: 0.015103660523891449
step: 500, loss: 0.07736295461654663
step: 510, loss: 0.16256630420684814
step: 520, loss: 0.002398537937551737
step: 530, loss: 0.02884935960173607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9238673517048108, f1=0.9232914923291492,

```



```
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10501015186309814
step: 10, loss: 0.02207767777144909
step: 20, loss: 0.0038011674769222736
step: 30, loss: 0.09509924054145813
step: 40, loss: 0.011400713585317135
step: 50, loss: 0.016228633001446724
step: 60, loss: 0.006177045404911041
step: 70, loss: 0.010879465378820896
step: 80, loss: 0.0035539199598133564
step: 90, loss: 0.0009259554208256304
step: 100, loss: 0.0048835184425115585
step: 110, loss: 0.002209028694778681
step: 120, loss: 0.01522222999483347
step: 130, loss: 0.0012930551311001182
step: 140, loss: 0.004833132028579712
step: 150, loss: 0.010881110094487667
step: 160, loss: 0.0030484984163194895
step: 170, loss: 0.02089509181678295
step: 180, loss: 0.17789921164512634
step: 190, loss: 0.0033127155620604753
step: 200, loss: 0.00973569881170988
step: 210, loss: 0.0019784148316830397
step: 220, loss: 0.004577151499688625
step: 230, loss: 0.0009747868170961738
step: 240, loss: 0.08039645850658417
step: 250, loss: 0.007913777604699135
step: 260, loss: 0.07740479707717896
step: 270, loss: 0.001622293726541102
step: 280, loss: 0.007190435193479061
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 290, loss: 0.20281237363815308
step: 300, loss: 0.0053489431738853455
step: 310, loss: 0.002891476033255458
step: 320, loss: 0.01316593587398529
step: 330, loss: 0.03278180956840515
step: 340, loss: 0.16700169444084167
step: 350, loss: 0.033616963773965836
step: 360, loss: 0.004128642380237579
step: 370, loss: 0.041178736835718155
step: 380, loss: 0.005752590950578451
step: 390, loss: 0.0051356153562664986
step: 400, loss: 0.28911492228507996
step: 410, loss: 0.035610735416412354
step: 420, loss: 0.002873928053304553
```

```
step: 430, loss: 0.008111400529742241
step: 440, loss: 0.007176954299211502
step: 450, loss: 0.013701566495001316
step: 460, loss: 0.005585273262113333
step: 470, loss: 0.14043648540973663
step: 480, loss: 0.005964913871139288
step: 490, loss: 0.08469241112470627
step: 500, loss: 0.019702095538377762
step: 510, loss: 0.0010489409323781729
step: 520, loss: 0.0017738559981808066
step: 530, loss: 0.0026496390346437693
/usr/local/lib/python3.7/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=0.9216043755697356, f1=0.9125683060109289,
best_f1=0.9232914923291492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004012977704405785
step: 10, loss: 0.0985226184129715
step: 20, loss: 0.015622572042047977
step: 30, loss: 0.025683004409074783
step: 40, loss: 0.006950364913791418
step: 50, loss: 0.005830494686961174
step: 60, loss: 0.07673219591379166
step: 70, loss: 0.0027386893052607775
step: 80, loss: 0.0015306684654206038
step: 90, loss: 0.009311988949775696
step: 100, loss: 0.025752047076821327
step: 110, loss: 0.0013349559158086777
step: 120, loss: 0.00319209904409945
step: 130, loss: 0.02932463213801384
step: 140, loss: 0.2258920967578888
step: 150, loss: 0.004268335644155741
step: 160, loss: 0.0028441878966987133
step: 170, loss: 0.05479341745376587
step: 180, loss: 0.0038359917234629393
step: 190, loss: 0.09724009782075882
step: 200, loss: 0.005253995768725872
step: 210, loss: 0.2047569304704666
step: 220, loss: 0.0015345580177381635
step: 230, loss: 0.05508190393447876
step: 240, loss: 0.040642403066158295
step: 250, loss: 0.003808635286986828
step: 260, loss: 0.0013783914037048817
step: 270, loss: 0.02492378279566765
step: 280, loss: 0.010311906225979328
```

```

step: 290, loss: 0.04919526353478432
step: 300, loss: 0.0017262650653719902
step: 310, loss: 0.015141579322516918
step: 320, loss: 0.0049286638386547565
step: 330, loss: 0.10132134705781937
step: 340, loss: 0.035868726670742035
step: 350, loss: 0.0007314517861232162
step: 360, loss: 0.022731022909283638
step: 370, loss: 0.11587131768465042
step: 380, loss: 0.0026036230847239494
step: 390, loss: 0.0033268628176301718
step: 400, loss: 0.009251156821846962
step: 410, loss: 0.03747495263814926
step: 420, loss: 0.006303034722805023
step: 430, loss: 0.0010071491124108434
step: 440, loss: 0.020210135728120804
step: 450, loss: 0.011574364267289639
step: 460, loss: 0.014356336556375027
step: 470, loss: 0.0790020301938057
step: 480, loss: 0.015409826301038265
step: 490, loss: 0.037560705095529556
step: 500, loss: 0.0010831007966771722
step: 510, loss: 0.1544981300830841
step: 520, loss: 0.001768853748217225
step: 530, loss: 0.009600410237908363
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9257722452743199, f1=0.9130434782608695,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015163797652348876
step: 10, loss: 0.01659390702843666
step: 20, loss: 0.016154251992702484
step: 30, loss: 0.07912939041852951
step: 40, loss: 0.002970188157632947
step: 50, loss: 0.02704489603638649
step: 60, loss: 0.002572939032688737
step: 70, loss: 0.010081132873892784
step: 80, loss: 0.18791097402572632
step: 90, loss: 0.0412900485098362
step: 100, loss: 0.0020864016842097044

```

```
step: 110, loss: 0.17503143846988678
step: 120, loss: 0.0507887601852417
step: 130, loss: 0.01187089178711176
step: 140, loss: 0.010608943179249763
step: 150, loss: 0.01396444346755743
step: 160, loss: 0.004205103497952223
step: 170, loss: 0.026084251701831818
step: 180, loss: 0.20834234356880188
step: 190, loss: 0.003672792576253414
step: 200, loss: 0.00840406771749258
step: 210, loss: 0.19714149832725525
step: 220, loss: 0.0021318045910447836
step: 230, loss: 0.030214346945285797
step: 240, loss: 0.03767986595630646
step: 250, loss: 0.04798346385359764
step: 260, loss: 0.0029067164286971092
step: 270, loss: 0.0034761414863169193
step: 280, loss: 0.0022663038689643145
step: 290, loss: 0.006923946551978588
step: 300, loss: 0.02205900475382805
step: 310, loss: 0.008356927894055843
step: 320, loss: 0.0017449064180254936
step: 330, loss: 0.012744028121232986
step: 340, loss: 0.1126837506890297
step: 350, loss: 0.020594701170921326
step: 360, loss: 0.0012399459956213832
step: 370, loss: 0.001360153197310865
step: 380, loss: 0.004158307332545519
step: 390, loss: 0.0015682230005040765
step: 400, loss: 0.15446846187114716
step: 410, loss: 0.008747043088078499
step: 420, loss: 0.0047064973041415215
step: 430, loss: 0.020939424633979797
step: 440, loss: 0.0012210042914375663
step: 450, loss: 0.018608756363391876
step: 460, loss: 0.0011678199516609311
step: 470, loss: 0.015468657948076725
step: 480, loss: 0.001060259761288762
step: 490, loss: 0.0025018120650202036
step: 500, loss: 0.004551149904727936
step: 510, loss: 0.013017961755394936
step: 520, loss: 0.18548667430877686
step: 530, loss: 0.004543748218566179
/usr/local/lib/python3.7/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=0.9205328433624255, f1=0.9186733303044072,
best_f1=0.9130434782608695
```

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

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

```
step: 0, loss: 0.0036567479837685823  
step: 10, loss: 0.002251212950795889  
step: 20, loss: 0.0015706279082223773  
step: 30, loss: 0.0023645805194973946  
step: 40, loss: 0.0009794748621061444  
step: 50, loss: 0.0007037596078589559  
step: 60, loss: 0.0014986684545874596  
step: 70, loss: 0.002694555092602968  
step: 80, loss: 0.03988974913954735  
step: 90, loss: 0.008473369292914867  
step: 100, loss: 0.03364599123597145  
step: 110, loss: 0.009783929213881493  
step: 120, loss: 0.0007856017909944057  
step: 130, loss: 0.10580138117074966  
step: 140, loss: 0.0037035555578768253  
step: 150, loss: 0.007887299172580242  
step: 160, loss: 0.10933302342891693  
step: 170, loss: 0.00986415520310402  
step: 180, loss: 0.0036022942513227463  
step: 190, loss: 0.0019499697955325246  
step: 200, loss: 0.001222729915753007  
step: 210, loss: 0.0019901571795344353  
step: 220, loss: 0.0047682481817901134  
step: 230, loss: 0.0007458811742253602  
step: 240, loss: 0.0016404452035203576  
step: 250, loss: 0.005635388195514679  
step: 260, loss: 0.015194381587207317  
step: 270, loss: 0.0008700048783794045  
step: 280, loss: 0.06652333587408066  
step: 290, loss: 0.0012459177523851395  
step: 300, loss: 0.03633175045251846  
step: 310, loss: 0.06548359990119934  
step: 320, loss: 0.012010066770017147  
step: 330, loss: 0.0026208420749753714  
step: 340, loss: 0.0016380222514271736  
step: 350, loss: 0.0006561008049175143  
step: 360, loss: 0.0003287906583864242  
step: 370, loss: 0.0011079973774030805  
step: 380, loss: 0.000591972260735929  
step: 390, loss: 0.0026066857390105724  
step: 400, loss: 0.006370049901306629  
step: 410, loss: 0.0018013442168012261  
step: 420, loss: 0.0011760811321437359  
step: 430, loss: 0.0007921199430711567  
step: 440, loss: 0.0038201971910893917
```

```

step: 450, loss: 0.06317632645368576
step: 460, loss: 0.0016202405095100403
step: 470, loss: 0.015186609700322151
step: 480, loss: 0.006342931184917688
step: 490, loss: 0.04835737496614456
step: 500, loss: 0.009976787492632866
step: 510, loss: 0.04686089977622032
step: 520, loss: 0.23728260397911072
step: 530, loss: 0.015155879780650139
/usr/local/lib/python3.7/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=0.9246823956442831, f1=0.9151157512482978,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004691645503044128
step: 10, loss: 0.0265961941331625
step: 20, loss: 0.030887421220541
step: 30, loss: 0.0017487769946455956
step: 40, loss: 0.0004988955333828926
step: 50, loss: 0.0016658138483762741
step: 60, loss: 0.0006401002174243331
step: 70, loss: 0.005013283807784319
step: 80, loss: 0.0005511431954801083
step: 90, loss: 0.000937897537369281
step: 100, loss: 0.0035937426146119833
step: 110, loss: 0.0004886435344815254
step: 120, loss: 0.0005697210435755551
step: 130, loss: 0.0004334079276304692
step: 140, loss: 0.0010330491932108998
step: 150, loss: 0.007776985410600901
step: 160, loss: 0.00041327139479108155
step: 170, loss: 0.0051018111407756805
step: 180, loss: 0.00630735419690609
step: 190, loss: 0.00563550041988492
step: 200, loss: 0.0005984751041978598
step: 210, loss: 0.005397078115493059
step: 220, loss: 0.0013402157928794622
step: 230, loss: 0.013451017439365387
step: 240, loss: 0.00254154484719038
step: 250, loss: 0.03395452722907066
step: 260, loss: 0.0036126456689089537
step: 270, loss: 0.007790683768689632
step: 280, loss: 0.0032086200080811977
step: 290, loss: 0.000360988691681996
step: 300, loss: 0.12921807169914246

```

```

step: 310, loss: 0.03421911969780922
step: 320, loss: 0.008983093313872814
step: 330, loss: 0.0006895131664350629
step: 340, loss: 0.029261719435453415
step: 350, loss: 0.020682338625192642
step: 360, loss: 0.009830488823354244
step: 370, loss: 0.0038620547857135534
step: 380, loss: 0.0014512869529426098
step: 390, loss: 0.0007721766596660018
step: 400, loss: 0.0032302450854331255
step: 410, loss: 0.010747838765382767
step: 420, loss: 0.0035302285104990005
step: 430, loss: 0.0018517569405958056
step: 440, loss: 0.0023376806639134884
step: 450, loss: 0.0014837830094620585
step: 460, loss: 0.003163191257044673
step: 470, loss: 0.001881306990981102
step: 480, loss: 0.0031347908079624176
step: 490, loss: 0.002632132265716791
step: 500, loss: 0.0016785489860922098
step: 510, loss: 0.002844605129212141
step: 520, loss: 0.001574427355080843
step: 530, loss: 0.004092956893146038
/usr/local/lib/python3.7/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=0.9189189189189189, f1=0.9143897996357012,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009150452679023147
step: 10, loss: 0.0014096691738814116
step: 20, loss: 0.17969825863838196
step: 30, loss: 0.0007243262371048331
step: 40, loss: 0.0014026880962774158
step: 50, loss: 0.0015134032582864165
step: 60, loss: 0.006840801797807217
step: 70, loss: 0.0008777836337685585
step: 80, loss: 0.0009846326429396868
step: 90, loss: 0.0034589588176459074
step: 100, loss: 0.002700434299185872
step: 110, loss: 0.0008381738443858922
step: 120, loss: 0.0014708969974890351
step: 130, loss: 0.0016707169124856591
step: 140, loss: 0.007902368903160095
step: 150, loss: 0.0006314184283837676
step: 160, loss: 0.05621441826224327

```

```

step: 170, loss: 0.0012318645603954792
step: 180, loss: 0.0008037746883928776
step: 190, loss: 0.0010862494818866253
step: 200, loss: 0.0010483120568096638
step: 210, loss: 0.19801536202430725
step: 220, loss: 0.001158258761279285
step: 230, loss: 0.012947685085237026
step: 240, loss: 0.0033683404326438904
step: 250, loss: 0.005209184717386961
step: 260, loss: 0.0015906081534922123
step: 270, loss: 0.0025895629078149796
step: 280, loss: 0.0031487918458878994
step: 290, loss: 0.0034981579519808292
step: 300, loss: 0.011741465888917446
step: 310, loss: 0.0025048665702342987
step: 320, loss: 0.002531621605157852
step: 330, loss: 0.003538500051945448
step: 340, loss: 0.0005084144650027156
step: 350, loss: 0.0003558180760592222
step: 360, loss: 0.0006113371346145868
step: 370, loss: 0.00029899488436058164
step: 380, loss: 0.0004632028576452285
step: 390, loss: 0.014787690714001656
step: 400, loss: 0.00037155457539483905
step: 410, loss: 0.04817451909184456
step: 420, loss: 0.0003002102894242853
step: 430, loss: 0.008048463612794876
step: 440, loss: 0.02615337073802948
step: 450, loss: 0.0007803601329214871
step: 460, loss: 0.0006218502530828118
step: 470, loss: 0.002761293202638626
step: 480, loss: 0.005948053672909737
step: 490, loss: 0.0045326584950089455
step: 500, loss: 0.0034872102551162243
step: 510, loss: 0.0025775611866265535
step: 520, loss: 0.0038979605305939913
step: 530, loss: 0.004244874697178602
/usr/local/lib/python3.7/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=0.9189189189189189, f1=0.9137697516930023,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011406268458813429
step: 10, loss: 0.005292908754199743
step: 20, loss: 0.009033527225255966

```


step: 30, loss: 0.0018019620329141617
step: 40, loss: 0.0023066415451467037
step: 50, loss: 0.008561034686863422
step: 60, loss: 0.000979105127044022
step: 70, loss: 0.003067555371671915
step: 80, loss: 0.047696564346551895
step: 90, loss: 0.009290033020079136
step: 100, loss: 0.0005921151023358107
step: 110, loss: 0.0030332414899021387
step: 120, loss: 0.0017853158060461283
step: 130, loss: 0.00031311833299696445
step: 140, loss: 0.000756744178943336
step: 150, loss: 0.001676490530371666
step: 160, loss: 0.19387243688106537
step: 170, loss: 0.018412642180919647
step: 180, loss: 0.0017423885874450207
step: 190, loss: 0.0010798489674925804
step: 200, loss: 0.0007955605397000909
step: 210, loss: 0.001884011086076498
step: 220, loss: 0.0006206681136973202
step: 230, loss: 0.003783195512369275
step: 240, loss: 0.0017603242304176092
step: 250, loss: 0.0049663023091852665
step: 260, loss: 0.000683958234731108
step: 270, loss: 0.0036724978126585484
step: 280, loss: 0.10866707563400269
step: 290, loss: 0.0007166583091020584
step: 300, loss: 0.002208842197433114
step: 310, loss: 0.0006809354526922107
step: 320, loss: 0.0005737656028941274
step: 330, loss: 0.0011745308293029666
step: 340, loss: 0.014402462169528008
step: 350, loss: 0.00046257703797891736
step: 360, loss: 0.11535166203975677
step: 370, loss: 0.0041632382199168205
step: 380, loss: 0.001633987994864583
step: 390, loss: 0.0007936495239846408
step: 400, loss: 0.0010804540943354368
step: 410, loss: 0.0008431979222223163
step: 420, loss: 0.001810406451113522
step: 430, loss: 0.0242938045412302
step: 440, loss: 0.0006058414001017809
step: 450, loss: 0.000641107268165797
step: 460, loss: 0.0011184219038113952
step: 470, loss: 0.001885362551547587
step: 480, loss: 0.0019416327122598886
step: 490, loss: 0.012223012745380402
step: 500, loss: 0.00026096386136487126

```

step: 510, loss: 0.0008804411045275629
step: 520, loss: 0.0006605254602618515
step: 530, loss: 0.0014459743397310376
/usr/local/lib/python3.7/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=0.9203539823008848, f1=0.9194444444444445,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020873630419373512
step: 10, loss: 0.0012613923754543066
step: 20, loss: 0.0004583942936733365
step: 30, loss: 0.0022074186708778143
step: 40, loss: 0.00046738641685806215
step: 50, loss: 0.000568790128454566
step: 60, loss: 0.0010613470803946257
step: 70, loss: 0.0007946642581373453
step: 80, loss: 0.0006122337654232979
step: 90, loss: 0.00033192537375725806
step: 100, loss: 0.0007361328462138772
step: 110, loss: 0.00044317528954707086
step: 120, loss: 0.0004352094838395715
step: 130, loss: 0.0003870263753924519
step: 140, loss: 0.004387493245303631
step: 150, loss: 0.00044863048242405057
step: 160, loss: 0.0013719775015488267
step: 170, loss: 0.0030535710975527763
step: 180, loss: 0.00042174156988039613
step: 190, loss: 0.0005055476212874055
step: 200, loss: 0.0006175029557198286
step: 210, loss: 0.0005698434542864561
step: 220, loss: 0.000480829447042197
step: 230, loss: 0.0004235242959111929
step: 240, loss: 0.00023546034935861826
step: 250, loss: 0.009064515121281147
step: 260, loss: 0.0004121811653021723
step: 270, loss: 0.0005295826122164726
step: 280, loss: 0.005964611191302538
step: 290, loss: 0.0007313134265132248
step: 300, loss: 0.0004595088539645076
step: 310, loss: 0.0098929638043046
step: 320, loss: 0.004491677042096853
step: 330, loss: 0.0013515790924429893
step: 340, loss: 0.002245261799544096
step: 350, loss: 0.00039360413211397827
step: 360, loss: 0.000439281458966434

```

```
step: 370, loss: 0.0005070788902230561
step: 380, loss: 0.0008012718753889203
step: 390, loss: 0.0023645306937396526
step: 400, loss: 0.021091675385832787
step: 410, loss: 0.00023155577946454287
step: 420, loss: 0.005995175801217556
step: 430, loss: 0.0004960451042279601
step: 440, loss: 0.022029360756278038
step: 450, loss: 0.00044227830949239433
step: 460, loss: 0.017770428210496902
step: 470, loss: 0.0006563760107383132
step: 480, loss: 0.0005035363137722015
step: 490, loss: 0.003887646598741412
step: 500, loss: 0.1983850598335266
step: 510, loss: 0.02826749160885811
step: 520, loss: 0.00026039427029900253
step: 530, loss: 0.0006570414989255369
/usr/local/lib/python3.7/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=0.9235757295044, f1=0.9205909510618652,
best_f1=0.9130434782608695
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000685620354488492
step: 10, loss: 0.0020957994274795055
step: 20, loss: 0.001452335505746305
step: 30, loss: 0.02140270173549652
step: 40, loss: 0.0005848715081810951
step: 50, loss: 0.0015297731151804328
step: 60, loss: 0.0008107298635877669
step: 70, loss: 0.0007561798556707799
step: 80, loss: 0.00045278004836291075
step: 90, loss: 0.000367650791304186
step: 100, loss: 0.0002122376172337681
step: 110, loss: 0.14868929982185364
step: 120, loss: 0.0005801466759294271
step: 130, loss: 0.0016229747561737895
step: 140, loss: 0.001531061832793057
step: 150, loss: 0.0007574281771667302
step: 160, loss: 0.0006967707886360586
step: 170, loss: 0.0002584295580163598
step: 180, loss: 0.002704012906178832
step: 190, loss: 0.009394517168402672
step: 200, loss: 0.0005220242310315371
step: 210, loss: 0.001082940143533051
step: 220, loss: 0.0005869147134944797
```

```

step: 230, loss: 0.0010986273409798741
step: 240, loss: 0.0034582328516989946
step: 250, loss: 0.0010584791889414191
step: 260, loss: 0.0003900118754245341
step: 270, loss: 0.0005982285365462303
step: 280, loss: 0.000511002610437572
step: 290, loss: 0.007243099622428417
step: 300, loss: 0.00036729214480146766
step: 310, loss: 0.041408102959394455
step: 320, loss: 0.0010249100159853697
step: 330, loss: 0.0013091614237055182
step: 340, loss: 0.004806044511497021
step: 350, loss: 0.0034369854256510735
step: 360, loss: 0.0005120699643157423
step: 370, loss: 0.0008455016650259495
step: 380, loss: 0.0008612275123596191
step: 390, loss: 0.000502693117596209
step: 400, loss: 0.0010986081324517727
step: 410, loss: 0.0019210614264011383
step: 420, loss: 0.0020384013187140226
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 430, loss: 0.00027446020976640284
step: 440, loss: 0.18959514796733856
step: 450, loss: 0.0008807061240077019
step: 460, loss: 0.00044351289398036897
step: 470, loss: 0.0004656514502130449
step: 480, loss: 0.0009682451491244137
step: 490, loss: 0.0006480729207396507
step: 500, loss: 0.0003596169117372483
step: 510, loss: 0.0005702492780983448
step: 520, loss: 0.0005674765561707318
step: 530, loss: 0.00035922107053920627
/usr/local/lib/python3.7/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=0.92243381328379, f1=0.920796665122742,
best_f1=0.9130434782608695
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:34, 164.41it/s]
load_f1 = 0.9213793103448276
real_f1 = 0.9170549860205033
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:29, 147.32it/s]

```

1.1.4 Amazon-Google - Running the matcher

```

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

```

```

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

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

```

rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
  UserWarning)
step: 0, loss: 0.5065819621086121
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.49555039405822754
step: 20, loss: 0.5027462244033813
step: 30, loss: 0.29124391078948975
step: 40, loss: 0.31876257061958313
step: 50, loss: 0.43278881907463074
step: 60, loss: 0.4955633282661438
step: 70, loss: 0.3042401373386383
step: 80, loss: 0.3719005286693573
step: 90, loss: 0.2751767039299011
step: 100, loss: 0.22545351088047028
step: 110, loss: 0.2822023034095764
step: 120, loss: 0.38511574268341064
step: 130, loss: 0.25948479771614075
step: 140, loss: 0.47530362010002136
step: 150, loss: 0.37561458349227905
step: 160, loss: 0.5102120041847229
step: 170, loss: 0.22757142782211304
step: 180, loss: 0.320095032453537
step: 190, loss: 0.4974554181098938
step: 200, loss: 0.3054679036140442
step: 210, loss: 0.3119015693664551
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.30555555555555556, f1=0.30561797752808983,
best_f1=0.30561797752808983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33213019371032715
step: 10, loss: 0.15780100226402283
step: 20, loss: 0.5023573637008667
step: 30, loss: 0.5443429946899414
step: 40, loss: 0.44649413228034973
step: 50, loss: 0.26675158739089966
step: 60, loss: 0.3253406286239624
step: 70, loss: 0.44081681966781616
step: 80, loss: 0.3076709806919098
step: 90, loss: 0.36771920323371887

```

```

step: 100, loss: 0.5077871084213257
step: 110, loss: 0.3715517222881317
step: 120, loss: 0.2281557321548462
step: 130, loss: 0.17745822668075562
step: 140, loss: 0.24348647892475128
step: 150, loss: 0.4239214062690735
step: 160, loss: 0.18244795501232147
step: 170, loss: 0.5681547522544861
step: 180, loss: 0.3221178948879242
step: 190, loss: 0.2925656735897064
step: 200, loss: 0.1702505350112915
step: 210, loss: 0.3171185553073883
/usr/local/lib/python3.7/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=0.1852731591448931, f1=0.18519984170953702,
best_f1=0.30561797752808983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25398433208465576
step: 10, loss: 0.26919251680374146
step: 20, loss: 0.4791133403778076
step: 30, loss: 0.2715141177177429
step: 40, loss: 0.45798879861831665
step: 50, loss: 0.48420822620391846
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.5333207249641418
step: 70, loss: 0.1938079446554184
step: 80, loss: 0.44735482335090637
step: 90, loss: 0.24294739961624146
step: 100, loss: 0.3450112044811249
step: 110, loss: 0.24099892377853394
step: 120, loss: 0.24695149064064026
step: 130, loss: 0.15687936544418335
step: 140, loss: 0.3959440588951111
step: 150, loss: 0.31796470284461975
step: 160, loss: 0.2071683257818222
step: 170, loss: 0.3816579580307007
step: 180, loss: 0.24725496768951416
step: 190, loss: 0.16550230979919434
step: 200, loss: 0.2555943429470062
step: 210, loss: 0.27832263708114624
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.30561797752808983

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2829914689064026
step: 10, loss: 0.2811829447746277
step: 20, loss: 0.3338456451892853
step: 30, loss: 0.2518210709095001
step: 40, loss: 0.25160592794418335
step: 50, loss: 0.24989177286624908
step: 60, loss: 0.4900056719779968
step: 70, loss: 0.2521796226501465
step: 80, loss: 0.24904510378837585
step: 90, loss: 0.45277848839759827
step: 100, loss: 0.38018468022346497
step: 110, loss: 0.574101448059082
step: 120, loss: 0.38805586099624634
step: 130, loss: 0.6942037343978882
step: 140, loss: 0.5077409148216248
step: 150, loss: 0.3803161680698395
step: 160, loss: 0.31201720237731934
step: 170, loss: 0.17997628450393677
step: 180, loss: 0.08676380664110184
step: 190, loss: 0.16596779227256775
step: 200, loss: 0.27736207842826843
step: 210, loss: 0.39292776584625244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.30561797752808983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3688429892063141
step: 10, loss: 0.3204520642757416
step: 20, loss: 0.3118830621242523
step: 30, loss: 0.23533868789672852
step: 40, loss: 0.43399038910865784
step: 50, loss: 0.3744618892669678
step: 60, loss: 0.38819563388824463
step: 70, loss: 0.22943098843097687
step: 80, loss: 0.4271084666252136
step: 90, loss: 0.45880040526390076
step: 100, loss: 0.24574096500873566
step: 110, loss: 0.16516192257404327
step: 120, loss: 0.24789780378341675
step: 130, loss: 0.2826957106590271
step: 140, loss: 0.5292657613754272

```

```

step: 150, loss: 0.34256768226623535
step: 160, loss: 0.2335994392633438
step: 170, loss: 0.36919403076171875
step: 180, loss: 0.24129554629325867
step: 190, loss: 0.5207574367523193
step: 200, loss: 0.4055531620979309
step: 210, loss: 0.21476352214813232
/usr/local/lib/python3.7/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=0.23902439024390246, f1=0.2473684210526316,
best_f1=0.30561797752808983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1333063691854477
step: 10, loss: 0.38769373297691345
step: 20, loss: 0.3081243634223938
step: 30, loss: 0.2625757157802582
step: 40, loss: 0.2783297002315521
step: 50, loss: 0.5030919909477234
step: 60, loss: 0.2252502143383026
step: 70, loss: 0.34827205538749695
step: 80, loss: 0.2244080752134323
step: 90, loss: 0.2564213275909424
step: 100, loss: 0.23919008672237396
step: 110, loss: 0.25458383560180664
step: 120, loss: 0.17664596438407898
step: 130, loss: 0.39128732681274414
step: 140, loss: 0.318776398897171
step: 150, loss: 0.13356110453605652
step: 160, loss: 0.14964479207992554
step: 170, loss: 0.26696595549583435
step: 180, loss: 0.2821234166622162
step: 190, loss: 0.1951409876346588
step: 200, loss: 0.29206565022468567
step: 210, loss: 0.44061926007270813
/usr/local/lib/python3.7/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=0.25386996904024767, f1=0.2990654205607477,
best_f1=0.30561797752808983
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45984742045402527
step: 10, loss: 0.2673366963863373
step: 20, loss: 0.2883480191230774

```

```

step: 30, loss: 0.223985493183136
step: 40, loss: 0.10177639871835709
step: 50, loss: 0.2546747028827667
step: 60, loss: 0.2331138551235199
step: 70, loss: 0.20482118427753448
step: 80, loss: 0.20412947237491608
step: 90, loss: 0.306633859872818
step: 100, loss: 0.2730729877948761
step: 110, loss: 0.291764497756958
step: 120, loss: 0.25013500452041626
step: 130, loss: 0.40603017807006836
step: 140, loss: 0.2395932674407959
step: 150, loss: 0.23244275152683258
step: 160, loss: 0.3835136294364929
step: 170, loss: 0.5696312189102173
step: 180, loss: 0.1781913936138153
step: 190, loss: 0.22178325057029724
step: 200, loss: 0.21187593042850494
step: 210, loss: 0.2995346784591675
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.49914529914529915, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30750706791877747
step: 10, loss: 0.20194511115550995
step: 20, loss: 0.16911791265010834
step: 30, loss: 0.34724050760269165
step: 40, loss: 0.1732526421546936
step: 50, loss: 0.08372803777456284
step: 60, loss: 0.13521891832351685
step: 70, loss: 0.3633383810520172
step: 80, loss: 0.2124808430671692
step: 90, loss: 0.3580867350101471
step: 100, loss: 0.23321110010147095
step: 110, loss: 0.18804334104061127
step: 120, loss: 0.2705872356891632
step: 130, loss: 0.05917583033442497
step: 140, loss: 0.23120525479316711
step: 150, loss: 0.3558219075202942
step: 160, loss: 0.23279419541358948
step: 170, loss: 0.30315932631492615

```

```

step: 180, loss: 0.24486394226551056
step: 190, loss: 0.12324802577495575
step: 200, loss: 0.1882697194814682
step: 210, loss: 0.2838694453239441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5205479452054794, f1=0.5308219178082191,
best_f1=0.5308219178082191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.264043390750885
step: 10, loss: 0.24249373376369476
step: 20, loss: 0.14704087376594543
step: 30, loss: 0.10017545521259308
step: 40, loss: 0.08500445634126663
step: 50, loss: 0.4289775788784027
step: 60, loss: 0.10483985394239426
step: 70, loss: 0.3085933327674866
step: 80, loss: 0.15304934978485107
step: 90, loss: 0.30083996057510376
step: 100, loss: 0.16598625481128693
step: 110, loss: 0.23102878034114838
step: 120, loss: 0.4273487329483032
step: 130, loss: 0.18230365216732025
step: 140, loss: 0.29915285110473633
step: 150, loss: 0.13024233281612396
step: 160, loss: 0.16693635284900665
step: 170, loss: 0.2088109850883484
step: 180, loss: 0.22826242446899414
step: 190, loss: 0.17900249361991882
step: 200, loss: 0.0854855552315712
step: 210, loss: 0.18814575672149658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5349794238683128, f1=0.5330578512396694,
best_f1=0.5330578512396694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2029745876789093
step: 10, loss: 0.18028314411640167
step: 20, loss: 0.16303081810474396
step: 30, loss: 0.035686250776052475
step: 40, loss: 0.16796109080314636
step: 50, loss: 0.18492422997951508
step: 60, loss: 0.05897346884012222
step: 70, loss: 0.12670700252056122
step: 80, loss: 0.0857561007142067
step: 90, loss: 0.11107026040554047
step: 100, loss: 0.2470582276582718
step: 110, loss: 0.05381181091070175
step: 120, loss: 0.2899174392223358
step: 130, loss: 0.02288675308227539
step: 140, loss: 0.0805114135146141
step: 150, loss: 0.07252947986125946
step: 160, loss: 0.10479419678449631
step: 170, loss: 0.044612426310777664
step: 180, loss: 0.09795096516609192
step: 190, loss: 0.08415444195270538
step: 200, loss: 0.28096774220466614
step: 210, loss: 0.11323501914739609
/usr/local/lib/python3.7/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=0.5318471337579618, f1=0.54281098546042,
best_f1=0.5330578512396694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1450858861207962
step: 10, loss: 0.08217116445302963
step: 20, loss: 0.02624645084142685
step: 30, loss: 0.03313416987657547
step: 40, loss: 0.12248758971691132
step: 50, loss: 0.17248652875423431
step: 60, loss: 0.11512826383113861
step: 70, loss: 0.04687308892607689
step: 80, loss: 0.25571075081825256
step: 90, loss: 0.29135552048683167
step: 100, loss: 0.34516221284866333
step: 110, loss: 0.11340992897748947
step: 120, loss: 0.08529014140367508
step: 130, loss: 0.024884231388568878
step: 140, loss: 0.04209569841623306
step: 150, loss: 0.24697698652744293

```

```

step: 160, loss: 0.010779780335724354
step: 170, loss: 0.07810559868812561
step: 180, loss: 0.06346392631530762
step: 190, loss: 0.20351332426071167
step: 200, loss: 0.06188318878412247
step: 210, loss: 0.20884555578231812
/usr/local/lib/python3.7/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=0.525, f1=0.5539033457249071, best_f1=0.5330578512396694
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04249080643057823
step: 10, loss: 0.18806129693984985
step: 20, loss: 0.3089653253555298
step: 30, loss: 0.07189373672008514
step: 40, loss: 0.02473921701312065
step: 50, loss: 0.2986468970775604
step: 60, loss: 0.028413791209459305
step: 70, loss: 0.14558196067810059
step: 80, loss: 0.08965042233467102
step: 90, loss: 0.24180255830287933
step: 100, loss: 0.013772975653409958
step: 110, loss: 0.17000170052051544
step: 120, loss: 0.01800685189664364
step: 130, loss: 0.29954224824905396
step: 140, loss: 0.1297782063484192
step: 150, loss: 0.022460419684648514
step: 160, loss: 0.047492824494838715
step: 170, loss: 0.14762920141220093
step: 180, loss: 0.08005271106958389
step: 190, loss: 0.01940738782286644
step: 200, loss: 0.009337091818451881
step: 210, loss: 0.0740518793463707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5354058721934368, f1=0.5317324185248713,
best_f1=0.5317324185248713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02492845430970192

```

```

step: 10, loss: 0.03039163164794445
step: 20, loss: 0.03930874168872833
step: 30, loss: 0.031709618866443634
step: 40, loss: 0.08282869309186935
step: 50, loss: 0.10373611748218536
step: 60, loss: 0.1241961345076561
step: 70, loss: 0.20110656321048737
step: 80, loss: 0.04200339689850807
step: 90, loss: 0.1846109926700592
step: 100, loss: 0.024959541857242584
step: 110, loss: 0.06342066824436188
step: 120, loss: 0.09828800708055496
step: 130, loss: 0.05554359406232834
step: 140, loss: 0.0754702091217041
step: 150, loss: 0.10248152911663055
step: 160, loss: 0.17237983644008636
step: 170, loss: 0.04363808408379555
step: 180, loss: 0.056336089968681335
step: 190, loss: 0.1914200782775879
step: 200, loss: 0.0815994143486023
step: 210, loss: 0.021527202799916267
/usr/local/lib/python3.7/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=0.5304659498207884, f1=0.5335753176043557,
best_f1=0.5317324185248713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02209438756108284
step: 10, loss: 0.06517577171325684
step: 20, loss: 0.13495831191539764
step: 30, loss: 0.030706321820616722
step: 40, loss: 0.023430529981851578
step: 50, loss: 0.028690006583929062
step: 60, loss: 0.05294524505734444
step: 70, loss: 0.030955733731389046
step: 80, loss: 0.047308895736932755
step: 90, loss: 0.015081712044775486
step: 100, loss: 0.07538734376430511
step: 110, loss: 0.09955113381147385
step: 120, loss: 0.11396002769470215
step: 130, loss: 0.06573940068483353
step: 140, loss: 0.06196126714348793
step: 150, loss: 0.11683379858732224
step: 160, loss: 0.06113802641630173
step: 170, loss: 0.032069429755210876
step: 180, loss: 0.020615197718143463

```

```

step: 190, loss: 0.05892046168446541
step: 200, loss: 0.03861774131655693
step: 210, loss: 0.012795483693480492
/usr/local/lib/python3.7/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=0.5194805194805194, f1=0.5418060200668896,
best_f1=0.5317324185248713
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04855883866548538
step: 10, loss: 0.034218717366456985
step: 20, loss: 0.159603551030159
step: 30, loss: 0.03114980272948742
step: 40, loss: 0.13628056645393372
step: 50, loss: 0.02140801027417183
step: 60, loss: 0.22241556644439697
step: 70, loss: 0.11336888372898102
step: 80, loss: 0.05861866846680641
step: 90, loss: 0.012606680393218994
step: 100, loss: 0.020165111869573593
step: 110, loss: 0.26237332820892334
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.07491306960582733
step: 130, loss: 0.02448955364525318
step: 140, loss: 0.09363394975662231
step: 150, loss: 0.050403252243995667
step: 160, loss: 0.17744304239749908
step: 170, loss: 0.031423795968294144
step: 180, loss: 0.033867593854665756
step: 190, loss: 0.04957297816872597
step: 200, loss: 0.02840203046798706
step: 210, loss: 0.10210806876420975
/usr/local/lib/python3.7/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=0.5272108843537415, f1=0.5342706502636204,
best_f1=0.5317324185248713
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                   : True
opt_level                 : 02
cast_model_type           : torch.float16
patch_torch_functions     : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:09, 231.18it/s]
load_f1 = 0.5250431778929189
real_f1 = 0.5217391304347826
267it [00:00, 1211.54it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 144.50it/s]

```

1.1.5 Walmart-Amazon - Running the matcher

```

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

```

```

--da span_del \
--summarize

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

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

```

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

Defaults for this optimization level are:

```

enabled          : True
opt_level        : 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: 0.4422927796840668
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.41409310698509216
step: 20, loss: 0.32544437050819397
step: 30, loss: 0.40906378626823425
step: 40, loss: 0.24761825799942017
step: 50, loss: 0.31688305735588074
step: 60, loss: 0.4866802990436554
step: 70, loss: 0.4059370756149292
step: 80, loss: 0.18210797011852264
step: 90, loss: 0.3153109550476074
step: 100, loss: 0.43451833724975586
step: 110, loss: 0.28668680787086487
step: 120, loss: 0.3464924395084381
step: 130, loss: 0.3472649157047272
step: 140, loss: 0.17474211752414703
step: 150, loss: 0.3111874461174011
step: 160, loss: 0.23043999075889587
step: 170, loss: 0.35315608978271484
step: 180, loss: 0.14174070954322815
step: 190, loss: 0.16247199475765228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.39999999999999997, f1=0.405458089668616,
best_f1=0.405458089668616
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.35986989736557007
step: 10, loss: 0.34668806195259094
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.7110645771026611
step: 30, loss: 0.24589204788208008
step: 40, loss: 0.31077539920806885
step: 50, loss: 0.4048060178756714
step: 60, loss: 0.42475640773773193
step: 70, loss: 0.3224136233329773
step: 80, loss: 0.19063366949558258
```

```

step: 90, loss: 0.29442620277404785
step: 100, loss: 0.24961301684379578
step: 110, loss: 0.37375056743621826
step: 120, loss: 0.23218320310115814
step: 130, loss: 0.46910005807876587
step: 140, loss: 0.3113897144794464
step: 150, loss: 0.3519757390022278
step: 160, loss: 0.2502190172672272
step: 170, loss: 0.19273990392684937
step: 180, loss: 0.15336520969867706
step: 190, loss: 0.10529036819934845
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4195121951219512, f1=0.460093896713615,
best_f1=0.460093896713615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4220387041568756
step: 10, loss: 0.29722389578819275
step: 20, loss: 0.27691972255706787
step: 30, loss: 0.17167600989341736
step: 40, loss: 0.09400293976068497
step: 50, loss: 0.24306637048721313
step: 60, loss: 0.13951432704925537
step: 70, loss: 0.25234079360961914
step: 80, loss: 0.14888954162597656
step: 90, loss: 0.08500877767801285
step: 100, loss: 0.1613772213459015
step: 110, loss: 0.4575992524623871
step: 120, loss: 0.29300248622894287
step: 130, loss: 0.09385479241609573
step: 140, loss: 0.057957254350185394
step: 150, loss: 0.4309006333351135
step: 160, loss: 0.20312298834323883
step: 170, loss: 0.48355454206466675
step: 180, loss: 0.18041406571865082
step: 190, loss: 0.1289614588022232
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7513227513227512, f1=0.7724867724867726,
best_f1=0.7724867724867726
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07140780985355377
step: 10, loss: 0.2034432291984558
step: 20, loss: 0.18512868881225586
step: 30, loss: 0.03261328488588333
step: 40, loss: 0.10274912416934967
step: 50, loss: 0.060377366840839386
step: 60, loss: 0.14623773097991943
step: 70, loss: 0.031918272376060486
step: 80, loss: 0.06516040116548538
step: 90, loss: 0.024382898584008217
step: 100, loss: 0.21663112938404083
step: 110, loss: 0.27006903290748596
step: 120, loss: 0.09533444792032242
step: 130, loss: 0.1639796644449234
step: 140, loss: 0.16995210945606232
step: 150, loss: 0.1607610434293747
step: 160, loss: 0.06679418683052063
step: 170, loss: 0.052404653280973434
step: 180, loss: 0.07460153102874756
step: 190, loss: 0.022547459229826927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8021390374331551, f1=0.8225806451612903,
best_f1=0.8225806451612903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15385565161705017
step: 10, loss: 0.1621655523777008
step: 20, loss: 0.01902157813310623
step: 30, loss: 0.07840847969055176
step: 40, loss: 0.09860822558403015
step: 50, loss: 0.06318444013595581
step: 60, loss: 0.004941849038004875
step: 70, loss: 0.24780358374118805
step: 80, loss: 0.013822120614349842

```

```

step: 90, loss: 0.06372038275003433
step: 100, loss: 0.036353904753923416
step: 110, loss: 0.1368672400712967
step: 120, loss: 0.24289143085479736
step: 130, loss: 0.19530470669269562
step: 140, loss: 0.0677742063999176
step: 150, loss: 0.11469512432813644
step: 160, loss: 0.06566904485225677
step: 170, loss: 0.0061753978952765465
step: 180, loss: 0.0555865578353405
step: 190, loss: 0.12352076917886734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8184281842818427, f1=0.8346883468834688,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07568302005529404
step: 10, loss: 0.04024016857147217
step: 20, loss: 0.07030551880598068
step: 30, loss: 0.12585316598415375
step: 40, loss: 0.06028800830245018
step: 50, loss: 0.22236433625221252
step: 60, loss: 0.05161904916167259
step: 70, loss: 0.07725097984075546
step: 80, loss: 0.14765025675296783
step: 90, loss: 0.12054220587015152
step: 100, loss: 0.14079606533050537
step: 110, loss: 0.004207131452858448
step: 120, loss: 0.011158633977174759
step: 130, loss: 0.07788905501365662
step: 140, loss: 0.0038105607964098454
step: 150, loss: 0.015883473679423332
step: 160, loss: 0.2287253588438034
step: 170, loss: 0.21728813648223877
step: 180, loss: 0.035189248621463776
step: 190, loss: 0.11514163017272949
/usr/local/lib/python3.7/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=0.8096514745308311, f1=0.8167539267015705,
best_f1=0.8346883468834688

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012467917054891586
step: 10, loss: 0.09524158388376236
step: 20, loss: 0.006351935677230358
step: 30, loss: 0.01908287964761257
step: 40, loss: 0.015899933874607086
step: 50, loss: 0.0018725156551226974
step: 60, loss: 0.18724325299263
step: 70, loss: 0.0056974864564836025
step: 80, loss: 0.005995457526296377
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.07211531698703766
step: 100, loss: 0.14356911182403564
step: 110, loss: 0.23787711560726166
step: 120, loss: 0.007029885426163673
step: 130, loss: 0.1295914500951767
step: 140, loss: 0.07789154350757599
step: 150, loss: 0.041053276509046555
step: 160, loss: 0.09590721130371094
step: 170, loss: 0.0838974267244339
step: 180, loss: 0.1171867847442627
step: 190, loss: 0.1642819195985794
/usr/local/lib/python3.7/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=0.7869674185463658, f1=0.8153846153846155,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10871706157922745
step: 10, loss: 0.02430316060781479
step: 20, loss: 0.01094183698296547
step: 30, loss: 0.16032782196998596
step: 40, loss: 0.05893300101161003
step: 50, loss: 0.04453577473759651
step: 60, loss: 0.043273985385894775
step: 70, loss: 0.017044885084033012
step: 80, loss: 0.20851218700408936
step: 90, loss: 0.047314856201410294
step: 100, loss: 0.011095613241195679
step: 110, loss: 0.21190455555915833
step: 120, loss: 0.0922723263502121
step: 130, loss: 0.013334646821022034
step: 140, loss: 0.01972431316971779
step: 150, loss: 0.0370931476354599

```

```

step: 160, loss: 0.001504286308772862
step: 170, loss: 0.21040870249271393
step: 180, loss: 0.03785346448421478
step: 190, loss: 0.04098917916417122
/usr/local/lib/python3.7/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=0.773067331670823, f1=0.7979539641943734,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004622331820428371
step: 10, loss: 0.0014756574528291821
step: 20, loss: 0.02787819877266884
step: 30, loss: 0.017538610845804214
step: 40, loss: 0.09332991391420364
step: 50, loss: 0.009741660207509995
step: 60, loss: 0.01520139817148447
step: 70, loss: 0.13682110607624054
step: 80, loss: 0.0031182230450212955
step: 90, loss: 0.06829529255628586
step: 100, loss: 0.005257747136056423
step: 110, loss: 0.030822718515992165
step: 120, loss: 0.002975975628942251
step: 130, loss: 0.07077882438898087
step: 140, loss: 0.206937775015831
step: 150, loss: 0.011869892477989197
step: 160, loss: 0.013742850162088871
step: 170, loss: 0.008577234111726284
step: 180, loss: 0.006597996223717928
step: 190, loss: 0.002933636773377657
/usr/local/lib/python3.7/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=0.7919799498746868, f1=0.8102564102564102,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0037466431967914104
step: 10, loss: 0.038845039904117584
step: 20, loss: 0.016598911955952644
step: 30, loss: 0.10876806825399399
step: 40, loss: 0.006768432445824146
step: 50, loss: 0.0020321072079241276
step: 60, loss: 0.0017500682733953
step: 70, loss: 0.0036165642086416483

```



```

step: 80, loss: 0.00907049234956503
step: 90, loss: 0.006749528460204601
step: 100, loss: 0.005654171574860811
step: 110, loss: 0.06147634983062744
step: 120, loss: 0.0045119295828044415
step: 130, loss: 0.0017243691254407167
step: 140, loss: 0.0025894527789205313
step: 150, loss: 0.014479683712124825
step: 160, loss: 0.0019843997433781624
step: 170, loss: 0.0008840336813591421
step: 180, loss: 0.006510570179671049
step: 190, loss: 0.001753023942001164
/usr/local/lib/python3.7/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=0.7967914438502673, f1=0.8266666666666667,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013846231624484062
step: 10, loss: 0.0007782697794027627
step: 20, loss: 0.0028479972388595343
step: 30, loss: 0.17238472402095795
step: 40, loss: 0.0023527462035417557
step: 50, loss: 0.001000315765850246
step: 60, loss: 0.004634345881640911
step: 70, loss: 0.02691614255309105
step: 80, loss: 0.07293510437011719
step: 90, loss: 0.0005022595287300646
step: 100, loss: 0.006856773514300585
step: 110, loss: 0.017416780814528465
step: 120, loss: 0.0015412714565172791
step: 130, loss: 0.0059960754588246346
step: 140, loss: 0.001860815566033125
step: 150, loss: 0.0038337286096066236
step: 160, loss: 0.0008916427032090724
step: 170, loss: 0.0008527648169547319
step: 180, loss: 0.18274015188217163
step: 190, loss: 0.00162572565022856
/usr/local/lib/python3.7/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=0.7989821882951654, f1=0.8082901554404145,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.021235773339867592
step: 10, loss: 0.0018341972026973963
step: 20, loss: 0.00047729327343404293
step: 30, loss: 0.02423781156539917
step: 40, loss: 0.0011873898329213262
step: 50, loss: 0.011660690419375896
step: 60, loss: 0.003914512228220701
step: 70, loss: 0.001386126852594316
step: 80, loss: 0.002265702001750469
step: 90, loss: 0.1496557742357254
step: 100, loss: 0.0014980320120230317
step: 110, loss: 0.013311675749719143
step: 120, loss: 0.0101112499833107
step: 130, loss: 0.010119706392288208
step: 140, loss: 0.0051437499932944775
step: 150, loss: 0.00896557979285717
step: 160, loss: 0.01920558698475361
step: 170, loss: 0.003471375908702612
step: 180, loss: 0.002873901277780533
step: 190, loss: 0.001517430180683732
/usr/local/lib/python3.7/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=0.7835616438356164, f1=0.8279569892473119,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020920580253005028
step: 10, loss: 0.021181579679250717
step: 20, loss: 0.009583834558725357
step: 30, loss: 0.09481120854616165
step: 40, loss: 0.0021839344408363104
step: 50, loss: 0.006871088873594999
step: 60, loss: 0.0011605279287323356
step: 70, loss: 0.000808277924079448
step: 80, loss: 0.000646025815512985
step: 90, loss: 0.0012245166581124067
step: 100, loss: 0.000573089171666652
step: 110, loss: 0.0152945127338171
step: 120, loss: 0.002188538433983922
step: 130, loss: 0.0045098052360117435
step: 140, loss: 0.00385781261138618
step: 150, loss: 0.05207080766558647
step: 160, loss: 0.036667317152023315
step: 170, loss: 0.00272530410438776
step: 180, loss: 0.007953735068440437
step: 190, loss: 0.04741779342293739

```

```

/usr/local/lib/python3.7/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=0.8085106382978724, f1=0.8213333333333334,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02418673224747181
step: 10, loss: 0.0011584277963265777
step: 20, loss: 0.0005950785125605762
step: 30, loss: 0.0007228407193906605
step: 40, loss: 0.007938901893794537
step: 50, loss: 0.0025574197061359882
step: 60, loss: 0.06156151369214058
step: 70, loss: 0.0017860213993117213
step: 80, loss: 0.0005719165783375502
step: 90, loss: 0.0018519412260502577
step: 100, loss: 0.004791425075381994
step: 110, loss: 0.014746167697012424
step: 120, loss: 0.03566556051373482
step: 130, loss: 0.0010636416263878345
step: 140, loss: 0.009127047844231129
step: 150, loss: 0.01326728519052267
step: 160, loss: 0.0006958436570130289
step: 170, loss: 0.0009233799064531922
step: 180, loss: 0.0012343230191618204
step: 190, loss: 0.002348574809730053
/usr/local/lib/python3.7/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=0.8085106382978724, f1=0.8213333333333334,
best_f1=0.8346883468834688
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011898830998688936
step: 10, loss: 0.0014948770403862
step: 20, loss: 0.03716828301548958
step: 30, loss: 0.0008018446387723088
step: 40, loss: 0.048574455082416534
step: 50, loss: 0.0022076584864407778
step: 60, loss: 0.0013282527215778828
step: 70, loss: 0.005780709441751242
step: 80, loss: 0.003941587172448635
step: 90, loss: 0.020133126527071
step: 100, loss: 0.0004997252253815532
step: 110, loss: 0.04008844122290611

```

```

step: 120, loss: 0.005750284995883703
step: 130, loss: 0.0012994854478165507
step: 140, loss: 0.0008093419019132853
step: 150, loss: 0.002237556502223015
step: 160, loss: 0.0010223676217719913
step: 170, loss: 0.016734635457396507
step: 180, loss: 0.0007122833631001413
step: 190, loss: 0.0012910082004964352
/usr/local/lib/python3.7/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=0.8106666666666666, f1=0.8213333333333334,
best_f1=0.8346883468834688
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:13, 150.75it/s]
load_f1 = 0.8102564102564102
real_f1 = 0.8
733it [00:00, 3294.47it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
4398it [00:30, 143.97it/s]
```

1.2 DA TEXTUAL

1.2.1 Abt-Buy - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                  : True
opt_level                : 02
cast_model_type          : torch.float16
patch_torch_functions    : False
keep_batchnorm_fp32      : True
master_weights           : True
loss_scale               : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.49235254526138306
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5098289251327515
step: 20, loss: 0.29689550399780273
step: 30, loss: 0.39665302634239197
step: 40, loss: 0.5697903037071228
step: 50, loss: 0.30461689829826355
step: 60, loss: 0.5367575287818909
step: 70, loss: 0.3520621657371521
step: 80, loss: 0.23777510225772858
step: 90, loss: 0.26740512251853943
step: 100, loss: 0.16569991409778595
step: 110, loss: 0.42603716254234314
step: 120, loss: 0.3442416787147522
step: 130, loss: 0.3018963038921356
step: 140, loss: 0.36068612337112427
step: 150, loss: 0.30384042859077454
step: 160, loss: 0.3982252776622772
step: 170, loss: 0.31325003504753113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20744081172491546, f1=0.2169222032935832,
best_f1=0.2169222032935832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2948252558708191
step: 10, loss: 0.46320536732673645
step: 20, loss: 0.31385859847068787
step: 30, loss: 0.2863704562187195
step: 40, loss: 0.08380362391471863
step: 50, loss: 0.3892049491405487
step: 60, loss: 0.19973202049732208
step: 70, loss: 0.470720112323761
step: 80, loss: 0.19658270478248596
step: 90, loss: 0.24735011160373688
step: 100, loss: 0.4814050793647766
step: 110, loss: 0.23632027208805084
step: 120, loss: 0.12785808742046356
step: 130, loss: 0.33999449014663696
step: 140, loss: 0.30530494451522827
step: 150, loss: 0.1816920042037964
step: 160, loss: 0.20218393206596375
step: 170, loss: 0.13093549013137817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7115384615384616, f1=0.6808510638297873,
best_f1=0.6808510638297873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41847485303878784
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.1848270148038864
step: 20, loss: 0.23522669076919556
step: 30, loss: 0.02429567277431488
step: 40, loss: 0.05595700442790985
step: 50, loss: 0.3433971703052521

```

```

step: 60, loss: 0.07859610766172409
step: 70, loss: 0.21774831414222717
step: 80, loss: 0.2217555046081543
step: 90, loss: 0.2660749554634094
step: 100, loss: 0.04817015677690506
step: 110, loss: 0.04534030333161354
step: 120, loss: 0.25696542859077454
step: 130, loss: 0.24972710013389587
step: 140, loss: 0.12346900254487991
step: 150, loss: 0.0915837287902832
step: 160, loss: 0.0651705339550972
step: 170, loss: 0.11424975097179413
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7727272727272727, f1=0.7121212121212122,
best_f1=0.7121212121212122
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06227083504199982
step: 10, loss: 0.2539524435997009
step: 20, loss: 0.052208539098501205
step: 30, loss: 0.27970120310783386
step: 40, loss: 0.09276780486106873
step: 50, loss: 0.08464747667312622
step: 60, loss: 0.26282984018325806
step: 70, loss: 0.02077663689851761
step: 80, loss: 0.4334898889064789
step: 90, loss: 0.22710749506950378
step: 100, loss: 0.1277080476284027
step: 110, loss: 0.2607320249080658
step: 120, loss: 0.4069218337535858
step: 130, loss: 0.1449524164199829
step: 140, loss: 0.18057143688201904
step: 150, loss: 0.21944858133792877
step: 160, loss: 0.028035620227456093
step: 170, loss: 0.027984196320176125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8088235294117646, f1=0.7570093457943925,
best_f1=0.7570093457943925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1112307459115982
step: 10, loss: 0.09245765954256058
step: 20, loss: 0.12726226449012756
step: 30, loss: 0.018103860318660736
step: 40, loss: 0.019678901880979538
step: 50, loss: 0.04801221191883087
step: 60, loss: 0.07169615477323532
step: 70, loss: 0.1294727623462677
step: 80, loss: 0.025788284838199615
step: 90, loss: 0.02719702199101448
step: 100, loss: 0.007815835066139698
step: 110, loss: 0.013477136380970478
step: 120, loss: 0.05185035988688469
step: 130, loss: 0.006297420710325241
step: 140, loss: 0.1221737340092659
step: 150, loss: 0.05755538493394852
step: 160, loss: 0.14785262942314148
step: 170, loss: 0.004289326258003712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8205128205128205, f1=0.7843137254901961,
best_f1=0.7843137254901961
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14761824905872345
step: 10, loss: 0.047271162271499634
step: 20, loss: 0.17530712485313416
step: 30, loss: 0.035412974655628204
step: 40, loss: 0.0021724626421928406
step: 50, loss: 0.01433294266462326
step: 60, loss: 0.022649817168712616
step: 70, loss: 0.02569042518734932
step: 80, loss: 0.024494873359799385
step: 90, loss: 0.09421252459287643
step: 100, loss: 0.02220999263226986
step: 110, loss: 0.04902447760105133

```

```

step: 120, loss: 0.010168012231588364
step: 130, loss: 0.19636861979961395
step: 140, loss: 0.13185663521289825
step: 150, loss: 0.15862837433815002
step: 160, loss: 0.01391275692731142
step: 170, loss: 0.006605606526136398
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8284313725490197, f1=0.7801418439716311,
best_f1=0.7801418439716311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016514429822564125
step: 10, loss: 0.021366072818636894
step: 20, loss: 0.01774064637720585
step: 30, loss: 0.3536234200000763
step: 40, loss: 0.00894930213689804
step: 50, loss: 0.0035356886219233274
step: 60, loss: 0.06433667242527008
step: 70, loss: 0.01078890636563301
step: 80, loss: 0.0074262917041778564
step: 90, loss: 0.01387022901326418
step: 100, loss: 0.0055985115468502045
step: 110, loss: 0.1387646496295929
step: 120, loss: 0.004272830672562122
step: 130, loss: 0.06882971525192261
step: 140, loss: 0.04064992070198059
step: 150, loss: 0.08564121276140213
step: 160, loss: 0.020920267328619957
step: 170, loss: 0.09326106309890747
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8325123152709358, f1=0.7754137115839244,
best_f1=0.7754137115839244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.02936648018658161
step: 10, loss: 0.046474892646074295
step: 20, loss: 0.032125674188137054
step: 30, loss: 0.0002490428159944713
step: 40, loss: 0.004474342800676823
step: 50, loss: 0.0432635098695755
step: 60, loss: 0.007080600131303072
step: 70, loss: 0.021608108654618263
step: 80, loss: 0.011001075617969036
step: 90, loss: 0.05072018876671791
step: 100, loss: 0.0007494384190067649
step: 110, loss: 0.09391949325799942
step: 120, loss: 0.002974834991618991
step: 130, loss: 0.003673334140330553
step: 140, loss: 0.023403828963637352
step: 150, loss: 0.00973766203969717
step: 160, loss: 0.02468654327094555
step: 170, loss: 0.08590700477361679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.834567901234568, f1=0.7895981087470448,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053671758621931076
step: 10, loss: 0.0038165391888469458
step: 20, loss: 0.005136832594871521
step: 30, loss: 0.06998128443956375
step: 40, loss: 0.023995842784643173
step: 50, loss: 0.0011700565228238702
step: 60, loss: 0.03506031632423401
step: 70, loss: 0.26223403215408325
step: 80, loss: 0.0021228373516350985
step: 90, loss: 0.01263468898832798
step: 100, loss: 0.045981355011463165
step: 110, loss: 0.013752954080700874
step: 120, loss: 0.0029198552947491407
step: 130, loss: 0.009915086440742016
step: 140, loss: 0.010303942486643791
step: 150, loss: 0.22934271395206451
step: 160, loss: 0.01718793623149395
step: 170, loss: 0.018628962337970734

```

```

/usr/local/lib/python3.7/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=0.802919708029197, f1=0.7647058823529412,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031855665147304535
step: 10, loss: 0.1262202262878418
step: 20, loss: 0.16969981789588928
step: 30, loss: 0.061377428472042084
step: 40, loss: 0.005908862687647343
step: 50, loss: 0.055816445499658585
step: 60, loss: 0.00839067529886961
step: 70, loss: 0.017213864251971245
step: 80, loss: 0.024231962859630585
step: 90, loss: 0.053947992622852325
step: 100, loss: 0.0014773006550967693
step: 110, loss: 0.008117358200252056
step: 120, loss: 0.0008015169878490269
step: 130, loss: 0.0016008717939257622
step: 140, loss: 0.0012454584939405322
step: 150, loss: 0.1652512550354004
step: 160, loss: 0.019276907667517662
step: 170, loss: 0.04283568263053894
/usr/local/lib/python3.7/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=0.8238213399503721, f1=0.7692307692307693,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06437047570943832
step: 10, loss: 0.007048470433801413
step: 20, loss: 0.0001925508986460045
step: 30, loss: 0.004177096765488386
step: 40, loss: 0.015721667557954788
step: 50, loss: 0.0016567904967814684
step: 60, loss: 0.006914779543876648
step: 70, loss: 0.004825785756111145
step: 80, loss: 0.12259458005428314
step: 90, loss: 0.008878654800355434
step: 100, loss: 0.022134924307465553
step: 110, loss: 0.009970585815608501
step: 120, loss: 0.0006582450005225837
step: 130, loss: 0.005372012034058571

```

```

step: 140, loss: 0.008172184228897095
step: 150, loss: 0.005258761812001467
step: 160, loss: 0.04881008341908455
step: 170, loss: 0.0021910294890403748
/usr/local/lib/python3.7/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=0.8256410256410256, f1=0.7990074441687345,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028251383919268847
step: 10, loss: 0.0006908312207087874
step: 20, loss: 0.003268349915742874
step: 30, loss: 0.03772815316915512
step: 40, loss: 0.010768122971057892
step: 50, loss: 0.004754107911139727
step: 60, loss: 0.0019108927808701992
step: 70, loss: 0.033018745481967926
step: 80, loss: 0.00015879000420682132
step: 90, loss: 0.012411192990839481
step: 100, loss: 0.0005580217111855745
step: 110, loss: 0.0025041031185537577
step: 120, loss: 0.03572998568415642
step: 130, loss: 0.0019361970480531454
step: 140, loss: 0.00014929278404451907
step: 150, loss: 0.00042022517300210893
step: 160, loss: 0.00036371522583067417
step: 170, loss: 0.0020858640782535076
/usr/local/lib/python3.7/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=0.8072916666666666, f1=0.7889447236180904,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0065927887335419655
step: 10, loss: 0.004416411276906729
step: 20, loss: 0.0004597783845383674
step: 30, loss: 0.008411243557929993
step: 40, loss: 0.001081477734260261
step: 50, loss: 0.02471449226140976
step: 60, loss: 0.001299968222156167
step: 70, loss: 0.07788458466529846
step: 80, loss: 0.00013141980161890388
step: 90, loss: 0.0005957631510682404

```

```

step: 100, loss: 0.03395865857601166
step: 110, loss: 0.005768731702119112
step: 120, loss: 0.005893555004149675
step: 130, loss: 0.00016175053315237164
step: 140, loss: 0.04277295991778374
step: 150, loss: 0.00027490119100548327
step: 160, loss: 0.006785823963582516
step: 170, loss: 0.00015763197734486312
/usr/local/lib/python3.7/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=0.8211586901763224, f1=0.7796610169491525,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028879977762699127
step: 10, loss: 0.0019367298809811473
step: 20, loss: 0.004484019707888365
step: 30, loss: 0.0042213331907987595
step: 40, loss: 0.00065524538513273
step: 50, loss: 0.00030734832398593426
step: 60, loss: 0.0019389602821320295
step: 70, loss: 0.006947537884116173
step: 80, loss: 0.019079996272921562
step: 90, loss: 0.000225709707592614
step: 100, loss: 0.007859482429921627
step: 110, loss: 0.004939975216984749
step: 120, loss: 0.001256493735127151
step: 130, loss: 0.061224017292261124
step: 140, loss: 0.018133951351046562
step: 150, loss: 0.016490748152136803
step: 160, loss: 0.0009042039746418595
step: 170, loss: 0.07551202923059464
/usr/local/lib/python3.7/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=0.8140703517587939, f1=0.7815533980582524,
best_f1=0.7895981087470448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010101227089762688
step: 10, loss: 0.0014237097930163145
step: 20, loss: 0.028051257133483887
step: 30, loss: 0.00042951753130182624
step: 40, loss: 0.0003905567864421755
step: 50, loss: 0.00014542760618496686

```

```

step: 60, loss: 0.00038055848563089967
step: 70, loss: 0.008486546576023102
step: 80, loss: 0.003254023613408208
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.007853282615542412
step: 100, loss: 0.008859360590577126
step: 110, loss: 0.00024549910449422896
step: 120, loss: 9.442191367270425e-05
step: 130, loss: 0.02507689781486988
step: 140, loss: 0.030768102034926414
step: 150, loss: 9.864967432804406e-05
step: 160, loss: 0.016578909009695053
step: 170, loss: 0.0001195978038595058
/usr/local/lib/python3.7/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=0.8089330024813897, f1=0.7761904761904763,
best_f1=0.7895981087470448
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

1916it [00:08, 231.52it/s]
load_f1 = 0.7405405405405405
real_f1 = 0.7335092348284961
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 145.06it/s]

```

1.3 DA DIRTY

1.3.1 DBLP-ACM - Running the matcher

```

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

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

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

```

```

Downloading: 100% 481/481 [00:00<00:00, 417kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 2.08MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 4.16MB/s]

```


Downloading: 100% 501M/501M [00:06<00:00, 72.2MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

step: 0, loss: 0.5850075483322144

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

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

step: 10, loss: 0.484156996011734

step: 20, loss: 0.5720878839492798

step: 30, loss: 0.2402663230895996

step: 40, loss: 0.35405758023262024

step: 50, loss: 0.5646131038665771

step: 60, loss: 0.44379445910453796

step: 70, loss: 0.2025277018547058

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

step: 80, loss: 0.28884145617485046

step: 90, loss: 0.2155524492263794

step: 100, loss: 0.1586422622203827

step: 110, loss: 0.19123078882694244

step: 120, loss: 0.217472106218338

step: 130, loss: 0.03961076959967613

step: 140, loss: 0.10698439180850983

step: 150, loss: 0.1279735416173935

step: 160, loss: 0.06075297296047211

step: 170, loss: 0.1217728704214096

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 180, loss: 0.22966298460960388
step: 190, loss: 0.017154064029455185
step: 200, loss: 0.10645948350429535
step: 210, loss: 0.04203534126281738
step: 220, loss: 0.11592482030391693
step: 230, loss: 0.03994778171181679
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9537444933920706, f1=0.9578713968957872,
best_f1=0.9578713968957872
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009035427123308182
step: 10, loss: 0.17645780742168427
step: 20, loss: 0.014681139960885048
step: 30, loss: 0.06525406241416931
step: 40, loss: 0.03980020433664322
step: 50, loss: 0.006776215508580208
step: 60, loss: 0.038745079189538956
step: 70, loss: 0.10968749970197678
step: 80, loss: 0.02718541771173477
step: 90, loss: 0.01826278306543827
step: 100, loss: 0.01655390113592148
step: 110, loss: 0.03709663078188896
step: 120, loss: 0.10244651883840561
step: 130, loss: 0.017832806333899498
step: 140, loss: 0.06511400640010834
step: 150, loss: 0.132195845246315
step: 160, loss: 0.0160610843449831
step: 170, loss: 0.0016882000491023064
step: 180, loss: 0.014063436537981033
step: 190, loss: 0.017840884625911713
step: 200, loss: 0.024665117263793945
step: 210, loss: 0.025970112532377243
step: 220, loss: 0.05258557200431824
step: 230, loss: 0.0014680515741929412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9582863585118377, f1=0.9517241379310345,
best_f1=0.9517241379310345
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012736041098833084
step: 10, loss: 0.012421488761901855
step: 20, loss: 0.07649245113134384
step: 30, loss: 0.05132708325982094
step: 40, loss: 0.006023343186825514
step: 50, loss: 0.020911717787384987
step: 60, loss: 0.018510911613702774
step: 70, loss: 0.1109013557434082
step: 80, loss: 0.0363314226269722
step: 90, loss: 0.005809234455227852
step: 100, loss: 0.014506781473755836
step: 110, loss: 0.043632201850414276
step: 120, loss: 0.004808720666915178
step: 130, loss: 0.006011276040226221
step: 140, loss: 0.0027150388341397047
step: 150, loss: 0.1116485670208931
step: 160, loss: 0.025192543864250183
step: 170, loss: 0.02152390405535698
step: 180, loss: 0.05534455180168152
step: 190, loss: 0.05548907816410065
step: 200, loss: 0.011921389028429985
step: 210, loss: 0.0027869022451341152
step: 220, loss: 0.07550419121980667
step: 230, loss: 0.03653300181031227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9662162162162162, 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: 0.023674657568335533
step: 10, loss: 0.01527270395308733
step: 20, loss: 0.0032508312724530697
step: 30, loss: 0.0016594667686149478
step: 40, loss: 0.07947637885808945

```

```

step: 50, loss: 0.012757294811308384
step: 60, loss: 0.02381238527595997
step: 70, loss: 0.11371422559022903
step: 80, loss: 0.11891257762908936
step: 90, loss: 0.09764615446329117
step: 100, loss: 0.013970177620649338
step: 110, loss: 0.004251571372151375
step: 120, loss: 0.020122574642300606
step: 130, loss: 0.02795739844441414
step: 140, loss: 0.0158880352973938
step: 150, loss: 0.008644734509289265
step: 160, loss: 0.0028054174035787582
step: 170, loss: 0.004306952003389597
step: 180, loss: 0.03503376245498657
step: 190, loss: 0.005496075842529535
step: 200, loss: 0.048899538815021515
step: 210, loss: 0.01410241611301899
step: 220, loss: 0.0009899038122966886
step: 230, loss: 0.001620929571799934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9692132269099202, f1=0.9502890173410403,
best_f1=0.9502890173410403
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00738714262843132
step: 10, loss: 0.023497579619288445
step: 20, loss: 0.02096370980143547
step: 30, loss: 0.004669730085879564
step: 40, loss: 0.010800698772072792
step: 50, loss: 0.02542324922978878
step: 60, loss: 0.018115615472197533
step: 70, loss: 0.0011941005941480398
step: 80, loss: 0.05051407217979431
step: 90, loss: 0.05465855076909065
step: 100, loss: 0.0005495318328030407
step: 110, loss: 0.0012089692754670978
step: 120, loss: 0.0016580312512814999
step: 130, loss: 0.006188681349158287
step: 140, loss: 0.04767724499106407
step: 150, loss: 0.022124402225017548
step: 160, loss: 0.005575126502662897

```

```

step: 170, loss: 0.003574963193386793
step: 180, loss: 0.03629503399133682
step: 190, loss: 0.009942063130438328
step: 200, loss: 0.008681409060955048
step: 210, loss: 0.015661969780921936
step: 220, loss: 0.005374344065785408
step: 230, loss: 0.10821506381034851
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9719416386083053, f1=0.9513023782559457,
best_f1=0.9513023782559457
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002273818012326956
step: 10, loss: 0.0008620147127658129
step: 20, loss: 0.0029615960083901882
step: 30, loss: 0.011437949724495411
step: 40, loss: 0.000817171239759773
step: 50, loss: 0.0006691793096251786
step: 60, loss: 0.006391379050910473
step: 70, loss: 0.17428207397460938
step: 80, loss: 0.011056025512516499
step: 90, loss: 0.001991491997614503
step: 100, loss: 0.0005913437926210463
step: 110, loss: 0.01251537911593914
step: 120, loss: 0.00057782611111117899
step: 130, loss: 0.006430257577449083
step: 140, loss: 0.0020592708606272936
step: 150, loss: 0.01267568115144968
step: 160, loss: 0.004222383722662926
step: 170, loss: 0.0004467711260076612
step: 180, loss: 0.004597908351570368
step: 190, loss: 0.001430472475476563
step: 200, loss: 0.008722356520593166
step: 210, loss: 0.0145291518419981
step: 220, loss: 0.015649553388357162
step: 230, loss: 0.001951441285200417
/usr/local/lib/python3.7/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=0.9711111111111111, f1=0.9654403567447045,
best_f1=0.9513023782559457

```

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

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

```
step: 0, loss: 0.002047439804300666  
step: 10, loss: 0.003530814778059721  
step: 20, loss: 0.0005020053358748555  
step: 30, loss: 0.00047516843187622726  
step: 40, loss: 0.0007454418810084462  
step: 50, loss: 0.000945655454415828  
step: 60, loss: 0.0013467719545587897  
step: 70, loss: 0.0012551106046885252  
step: 80, loss: 0.003911688458174467  
step: 90, loss: 0.0009887408232316375  
step: 100, loss: 0.0007380563183687627  
step: 110, loss: 0.0032440139912068844  
step: 120, loss: 0.0031551397405564785  
step: 130, loss: 0.005049690138548613  
step: 140, loss: 0.0007850852562114596  
step: 150, loss: 0.03958795219659805  
step: 160, loss: 0.0006361969863064587  
step: 170, loss: 0.012093433178961277  
step: 180, loss: 0.007171406410634518  
step: 190, loss: 0.02448847144842148  
step: 200, loss: 0.0015042823506519198  
step: 210, loss: 0.013813537545502186  
step: 220, loss: 0.0006820220150984824  
step: 230, loss: 0.0006742729456163943
```

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

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

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

```
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=0.9772727272727272, f1=0.9611872146118722,  
best_f1=0.9611872146118722
```

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

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

```
step: 0, loss: 0.0006521005416288972  
step: 10, loss: 0.0012788008898496628  
step: 20, loss: 0.0009307658183388412  
step: 30, loss: 0.003737750696018338  
step: 40, loss: 0.005400712136179209  
step: 50, loss: 0.004040514584630728  
step: 60, loss: 0.0010678237304091454  
step: 70, loss: 0.00026963985874317586  
step: 80, loss: 0.025013018399477005
```

```

step: 90, loss: 0.0006234640022739768
step: 100, loss: 0.0006887138588353992
step: 110, loss: 0.002689995104447007
step: 120, loss: 0.0006460452568717301
step: 130, loss: 0.03605359047651291
step: 140, loss: 0.0005976286483928561
step: 150, loss: 0.16787469387054443
step: 160, loss: 0.0022724969312548637
step: 170, loss: 0.055069632828235626
step: 180, loss: 0.0036548341158777475
step: 190, loss: 0.011985325254499912
step: 200, loss: 0.13327530026435852
step: 210, loss: 0.002752432134002447
step: 220, loss: 0.0012432645307853818
step: 230, loss: 0.0007777552818879485
/usr/local/lib/python3.7/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=0.9752808988764046, f1=0.963963963963964,
best_f1=0.9611872146118722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005061274860054255
step: 10, loss: 0.0006863146554678679
step: 20, loss: 0.001231155707500875
step: 30, loss: 0.0014811073197051883
step: 40, loss: 0.002060921862721443
step: 50, loss: 0.01829294115304947
step: 60, loss: 0.0007117573404684663
step: 70, loss: 0.018163669854402542
step: 80, loss: 0.003492768155410886
step: 90, loss: 0.16351047158241272
step: 100, loss: 0.00030104792676866055
step: 110, loss: 0.0003183303924743086
step: 120, loss: 0.07799991220235825
step: 130, loss: 0.011628732085227966
step: 140, loss: 0.0004974417388439178
step: 150, loss: 0.0002412226895103231
step: 160, loss: 0.0008629766525700688
step: 170, loss: 0.0002456303918734193
step: 180, loss: 0.00396139407530427
step: 190, loss: 8.918378443922848e-05
step: 200, loss: 0.0006211867439560592
step: 210, loss: 0.0008217755821533501
step: 220, loss: 0.01565036177635193
step: 230, loss: 0.00021160666074138135
/usr/local/lib/python3.7/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=0.9741863075196409, f1=0.9638009049773756,
best_f1=0.9611872146118722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00036116319824941456
step: 10, loss: 0.000607621215749532
step: 20, loss: 0.005182304419577122
step: 30, loss: 0.00012311873433645815
step: 40, loss: 0.00792614370584488
step: 50, loss: 0.00016351982776541263
step: 60, loss: 0.00020231102826073766
step: 70, loss: 0.005731641314923763
step: 80, loss: 0.004548089113086462
step: 90, loss: 0.0005733675789088011
step: 100, loss: 0.001001800294034183
step: 110, loss: 0.0030825070571154356
step: 120, loss: 0.0005982252769172192
step: 130, loss: 0.0007704953895881772
step: 140, loss: 0.00020620286522898823
step: 150, loss: 0.005786876659840345
step: 160, loss: 9.188867988996208e-05
step: 170, loss: 0.0001840086915763095
step: 180, loss: 0.004947856534272432
step: 190, loss: 0.0007262722938321531
step: 200, loss: 0.0009321649558842182
step: 210, loss: 0.006013920530676842
step: 220, loss: 0.0002166048507206142
step: 230, loss: 0.00021357531659305096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9807474518686297, f1=0.963718820861678,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017555431986693293
step: 10, loss: 0.0006842633592896163
step: 20, loss: 0.00023457813949789852
step: 30, loss: 0.00016504299128428102
step: 40, loss: 7.428186654578894e-05

```



```

step: 50, loss: 0.00021970612579025328
step: 60, loss: 0.01085524633526802
step: 70, loss: 0.00023215208784677088
step: 80, loss: 0.0005101553397253156
step: 90, loss: 0.19367016851902008
step: 100, loss: 0.0004114776966162026
step: 110, loss: 0.0054990206845104694
step: 120, loss: 0.0005309933912940323
step: 130, loss: 0.00027011078782379627
step: 140, loss: 0.0013312968658283353
step: 150, loss: 0.0004413897986523807
step: 160, loss: 0.0007705691386945546
step: 170, loss: 0.00758472690358758
step: 180, loss: 0.0004602769622579217
step: 190, loss: 0.00038238344131968915
step: 200, loss: 0.0015542850596830249
step: 210, loss: 0.00023946484725456685
step: 220, loss: 0.00031835745903663337
step: 230, loss: 0.0004929449642077088
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.9683972911963882,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004643940774258226
step: 10, loss: 0.00026026423438452184
step: 20, loss: 0.001981467241421342
step: 30, loss: 0.007200699765235186
step: 40, loss: 0.0002239397872472182
step: 50, loss: 0.0021689871791750193
step: 60, loss: 0.0005861124373041093
step: 70, loss: 0.00019211611652281135
step: 80, loss: 6.793371721869335e-05
step: 90, loss: 0.0011578060220927
step: 100, loss: 8.080551924649626e-05
step: 110, loss: 8.709450776223093e-05
step: 120, loss: 0.0002590655058156699
step: 130, loss: 0.00022967855329625309
step: 140, loss: 0.0006959949969314039
step: 150, loss: 0.00033293821616098285
step: 160, loss: 0.006475898902863264
step: 170, loss: 0.00035406355164013803
step: 180, loss: 0.0015987132210284472
step: 190, loss: 0.00037620370858348906
step: 200, loss: 0.0004955226904712617

```

```

step: 210, loss: 0.00314123323187232
step: 220, loss: 0.03420831635594368
step: 230, loss: 0.00033711371361278
/usr/local/lib/python3.7/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=0.9785794813979707, f1=0.9685393258426966,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005327956168912351
step: 10, loss: 0.0004490526916924864
step: 20, loss: 0.00044629626790992916
step: 30, loss: 0.00024252616276498884
step: 40, loss: 0.0008366873953491449
step: 50, loss: 0.003525296924635768
step: 60, loss: 0.0005832270835526288
step: 70, loss: 0.0005956381210125983
step: 80, loss: 0.006309507880359888
step: 90, loss: 0.00023212558880914003
step: 100, loss: 0.0005467826267704368
step: 110, loss: 0.0005579552380368114
step: 120, loss: 0.0002998782729264349
step: 130, loss: 0.0005510238115675747
step: 140, loss: 8.759072079556063e-05
step: 150, loss: 0.00019740939023904502
step: 160, loss: 0.0008912305347621441
step: 170, loss: 0.0002889930037781596
step: 180, loss: 0.032950326800346375
step: 190, loss: 0.0002837429929058999
step: 200, loss: 3.6557557905325666e-05
step: 210, loss: 0.0010231365449726582
step: 220, loss: 0.00041298248106613755
step: 230, loss: 0.00038595549995079637
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9652076318742986,
best_f1=0.9652076318742986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001764976914273575

```

```

step: 10, loss: 0.0002117412514053285
step: 20, loss: 0.0003892107051797211
step: 30, loss: 0.0013447507517412305
step: 40, loss: 0.00018194603035226464
step: 50, loss: 6.0781843785662204e-05
step: 60, loss: 0.0001556318165967241
step: 70, loss: 0.0002099411067320034
step: 80, loss: 0.000286593014607206
step: 90, loss: 0.0037779826670885086
step: 100, loss: 0.0017054929630830884
step: 110, loss: 0.0009010019130073488
step: 120, loss: 6.614124140469357e-05
step: 130, loss: 0.0005115754320286214
step: 140, loss: 0.0009047744679264724
step: 150, loss: 0.00047842503408901393
step: 160, loss: 0.0002619950973894447
step: 170, loss: 0.00013377550931181759
step: 180, loss: 0.010627007111907005
step: 190, loss: 0.00019120072829537094
step: 200, loss: 0.0006157051539048553
step: 210, loss: 0.0002469994651619345
step: 220, loss: 0.0014003097312524915
step: 230, loss: 0.00025665474822744727
/usr/local/lib/python3.7/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=0.9764309764309763, f1=0.9675977653631285,
best_f1=0.9652076318742986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04352060332894325
step: 10, loss: 0.0002381288359174505
step: 20, loss: 0.0014263620832934976
step: 30, loss: 0.0008856107597239316
step: 40, loss: 0.00029007942066527903
step: 50, loss: 0.00025933151482604444
step: 60, loss: 0.008786112070083618
step: 70, loss: 0.0016709580086171627
step: 80, loss: 0.00040850977529771626
step: 90, loss: 0.00021402222046162933
step: 100, loss: 0.00012027379852952436
step: 110, loss: 0.00010958637722069398
step: 120, loss: 0.05864306166768074
step: 130, loss: 0.00010806990030687302
step: 140, loss: 0.0102815181016922
step: 150, loss: 0.00012744627019856125
step: 160, loss: 0.0007813902338966727

```

```

step: 170, loss: 8.139592682709917e-05
step: 180, loss: 0.00029721984174102545
step: 190, loss: 0.0038179182447493076
step: 200, loss: 0.0025678849779069424
step: 210, loss: 0.0009580273763276637
step: 220, loss: 0.00019082137441728264
step: 230, loss: 0.0005557229742407799
/usr/local/lib/python3.7/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=0.976271186440678, f1=0.9660633484162895,
best_f1=0.9652076318742986
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 196.02it/s]
load_f1 = 0.9820224719101124
real_f1 = 0.9765363128491621
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:23, 183.79it/s]

1.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale              : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6021271347999573
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.48927000164985657
step: 20, loss: 0.28137579560279846
step: 30, loss: 0.3575286865234375
step: 40, loss: 0.32774513959884644
step: 50, loss: 0.4312056303024292
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.2407153993844986
step: 70, loss: 0.30559852719306946
step: 80, loss: 0.217124342918396
step: 90, loss: 0.21193501353263855
step: 100, loss: 0.2576051950454712
step: 110, loss: 0.2475411742925644
step: 120, loss: 0.15867742896080017
step: 130, loss: 0.12336186319589615
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.2521691918373108
step: 150, loss: 0.24412395060062408
step: 160, loss: 0.13285748660564423
step: 170, loss: 0.18323275446891785
step: 180, loss: 0.20185911655426025
step: 190, loss: 0.18593519926071167
step: 200, loss: 0.16255688667297363
step: 210, loss: 0.05090378224849701
step: 220, loss: 0.03489828482270241
step: 230, loss: 0.30177515745162964
step: 240, loss: 0.05433971807360649
step: 250, loss: 0.0330197848379612
step: 260, loss: 0.2743716537952423
step: 270, loss: 0.35621941089630127
step: 280, loss: 0.05251535028219223

```

```

step: 290, loss: 0.0627850741147995
step: 300, loss: 0.1559806913137436
step: 310, loss: 0.1465369015932083
step: 320, loss: 0.15394383668899536
step: 330, loss: 0.1402025818824768
step: 340, loss: 0.3821689188480377
step: 350, loss: 0.20290139317512512
step: 360, loss: 0.03662355616688728
step: 370, loss: 0.09656445682048798
step: 380, loss: 0.15540984272956848
step: 390, loss: 0.10217973589897156
step: 400, loss: 0.04591572284698486
step: 410, loss: 0.278754860162735
step: 420, loss: 0.021489422768354416
step: 430, loss: 0.041315607726573944
step: 440, loss: 0.030172375962138176
step: 450, loss: 0.18416619300842285
step: 460, loss: 0.08546731621026993
step: 470, loss: 0.04200410842895508
step: 480, loss: 0.11571329832077026
step: 490, loss: 0.14761732518672943
step: 500, loss: 0.1178479939699173
step: 510, loss: 0.2040894329547882
step: 520, loss: 0.38892021775245667
step: 530, loss: 0.09004048258066177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9041474654377881, f1=0.9005524861878452,
best_f1=0.9005524861878452
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10381624847650528
step: 10, loss: 0.14179623126983643
step: 20, loss: 0.1949073225259781
step: 30, loss: 0.059853650629520416
step: 40, loss: 0.19580356776714325
step: 50, loss: 0.2509353756904602
step: 60, loss: 0.03562203049659729
step: 70, loss: 0.0410626120865345
step: 80, loss: 0.1461990773677826
step: 90, loss: 0.055674728006124496
step: 100, loss: 0.15537574887275696

```

```
step: 110, loss: 0.0392434224486351
step: 120, loss: 0.09558281302452087
step: 130, loss: 0.02751762978732586
step: 140, loss: 0.0791824534535408
step: 150, loss: 0.053813111037015915
step: 160, loss: 0.018685828894376755
step: 170, loss: 0.2128654569387436
step: 180, loss: 0.19536761939525604
step: 190, loss: 0.057176701724529266
step: 200, loss: 0.22321532666683197
step: 210, loss: 0.0357600674033165
step: 220, loss: 0.006225903052836657
step: 230, loss: 0.12016916275024414
step: 240, loss: 0.12966719269752502
step: 250, loss: 0.0843336433172226
step: 260, loss: 0.08021166175603867
step: 270, loss: 0.03307768702507019
step: 280, loss: 0.039629630744457245
step: 290, loss: 0.042494628578424454
step: 300, loss: 0.15113162994384766
step: 310, loss: 0.041911520063877106
step: 320, loss: 0.18417659401893616
step: 330, loss: 0.09656403958797455
step: 340, loss: 0.14465685188770294
step: 350, loss: 0.006920136045664549
step: 360, loss: 0.07558457553386688
step: 370, loss: 0.06431883573532104
step: 380, loss: 0.17166177928447723
step: 390, loss: 0.0216812901198864
step: 400, loss: 0.12795014679431915
step: 410, loss: 0.19118836522102356
step: 420, loss: 0.08130518347024918
step: 430, loss: 0.101955346763134
step: 440, loss: 0.015658140182495117
step: 450, loss: 0.11181028187274933
step: 460, loss: 0.10854510217905045
step: 470, loss: 0.09498029947280884
step: 480, loss: 0.02972794696688652
step: 490, loss: 0.04405564069747925
step: 500, loss: 0.0227005984634161
step: 510, loss: 0.19244424998760223
step: 520, loss: 0.3794453740119934
step: 530, loss: 0.10961294919252396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.929368029739777, f1=0.9230769230769231,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13766154646873474
step: 10, loss: 0.23226284980773926
step: 20, loss: 0.017370279878377914
step: 30, loss: 0.06101859733462334
step: 40, loss: 0.07651188224554062
step: 50, loss: 0.12042830884456635
step: 60, loss: 0.03611094132065773
step: 70, loss: 0.039589375257492065
step: 80, loss: 0.1483336091041565
step: 90, loss: 0.0323101170361042
step: 100, loss: 0.15278975665569305
step: 110, loss: 0.10441923141479492
step: 120, loss: 0.1408575028181076
step: 130, loss: 0.04817638546228409
step: 140, loss: 0.024588342756032944
step: 150, loss: 0.04674936830997467
step: 160, loss: 0.037655871361494064
step: 170, loss: 0.02514350228011608
step: 180, loss: 0.03278575837612152
step: 190, loss: 0.03935021907091141
step: 200, loss: 0.046699728816747665
step: 210, loss: 0.06856601685285568
step: 220, loss: 0.11200110614299774
step: 230, loss: 0.06406103819608688
step: 240, loss: 0.21842694282531738
step: 250, loss: 0.1319742351770401
step: 260, loss: 0.11526448279619217
step: 270, loss: 0.04213925451040268
step: 280, loss: 0.051270030438899994
step: 290, loss: 0.00315583567135036
step: 300, loss: 0.3229296803474426
step: 310, loss: 0.05423345789313316
step: 320, loss: 0.08636312186717987
step: 330, loss: 0.004904761910438538
step: 340, loss: 0.02581176906824112
step: 350, loss: 0.1068231463432312
step: 360, loss: 0.017048027366399765
step: 370, loss: 0.047542523592710495
step: 380, loss: 0.008642377331852913
step: 390, loss: 0.03711691126227379
step: 400, loss: 0.18638063967227936

```

```

step: 410, loss: 0.027552513405680656
step: 420, loss: 0.022535935044288635
step: 430, loss: 0.11152851581573486
step: 440, loss: 0.3678845763206482
step: 450, loss: 0.03414585068821907
step: 460, loss: 0.1742609739303589
step: 470, loss: 0.00974488165229559
step: 480, loss: 0.1804620623588562
step: 490, loss: 0.10786823183298111
step: 500, loss: 0.14911320805549622
step: 510, loss: 0.08688095957040787
step: 520, loss: 0.003099879715591669
step: 530, loss: 0.03239227831363678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.929784304726939, f1=0.9144434222631096,
best_f1=0.9144434222631096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029294800013303757
step: 10, loss: 0.008641844615340233
step: 20, loss: 0.18510885536670685
step: 30, loss: 0.06187906488776207
step: 40, loss: 0.033923353999853134
step: 50, loss: 0.02228126861155033
step: 60, loss: 0.01766025274991989
step: 70, loss: 0.09794598072767258
step: 80, loss: 0.14240875840187073
step: 90, loss: 0.07564151287078857
step: 100, loss: 0.0028763869777321815
step: 110, loss: 0.14325737953186035
step: 120, loss: 0.0038432118017226458
step: 130, loss: 0.01304568350315094
step: 140, loss: 0.053687684237957
step: 150, loss: 0.014715554192662239
step: 160, loss: 0.019038410857319832
step: 170, loss: 0.01066505629569292
step: 180, loss: 0.08490083366632462
step: 190, loss: 0.0449758842587471
step: 200, loss: 0.037272512912750244
step: 210, loss: 0.0029867347329854965
step: 220, loss: 0.016164930537343025

```

```

step: 230, loss: 0.04373384267091751
step: 240, loss: 0.006161415483802557
step: 250, loss: 0.08538873493671417
step: 260, loss: 0.0057889665476977825
step: 270, loss: 0.02113775908946991
step: 280, loss: 0.010244578123092651
step: 290, loss: 0.1684872955083847
step: 300, loss: 0.008358447812497616
step: 310, loss: 0.008144109509885311
step: 320, loss: 0.006458647549152374
step: 330, loss: 0.020441723987460136
step: 340, loss: 0.06920942664146423
step: 350, loss: 0.026912806555628777
step: 360, loss: 0.1063561886548996
step: 370, loss: 0.005702035501599312
step: 380, loss: 0.0776008814573288
step: 390, loss: 0.005586352664977312
step: 400, loss: 0.021316256374120712
step: 410, loss: 0.002851000754162669
step: 420, loss: 0.016986966133117676
step: 430, loss: 0.0019328331109136343
step: 440, loss: 0.004689862951636314
step: 450, loss: 0.11684533208608627
step: 460, loss: 0.09055396914482117
step: 470, loss: 0.002395639894530177
step: 480, loss: 0.025210272520780563
step: 490, loss: 0.0013335166731849313
step: 500, loss: 0.056699465960264206
step: 510, loss: 0.058556728065013885
step: 520, loss: 0.01882915571331978
step: 530, loss: 0.1708500236272812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9372997711670481, f1=0.9288664525011474,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016921034548431635
step: 10, loss: 0.014089605771005154
step: 20, loss: 0.008025983348488808
step: 30, loss: 0.013129192404448986
step: 40, loss: 0.05321604013442993

```

step: 50, loss: 0.045059531927108765
step: 60, loss: 0.014983884058892727
step: 70, loss: 0.06685131043195724
step: 80, loss: 0.0009884759783744812
step: 90, loss: 0.16580845415592194
step: 100, loss: 0.015609601512551308
step: 110, loss: 0.056010618805885315
step: 120, loss: 0.12020819634199142
step: 130, loss: 0.009086444042623043
step: 140, loss: 0.00521048391237855
step: 150, loss: 0.010167153552174568
step: 160, loss: 0.02282230369746685
step: 170, loss: 0.01825261116027832
step: 180, loss: 0.002138679614290595
step: 190, loss: 0.01257923524826765
step: 200, loss: 0.035842448472976685
step: 210, loss: 0.005132411140948534
step: 220, loss: 0.03918416425585747
step: 230, loss: 0.04311414435505867
step: 240, loss: 0.044667601585388184
step: 250, loss: 0.09206278622150421
step: 260, loss: 0.001610989449545741
step: 270, loss: 0.012542682699859142
step: 280, loss: 0.00905867200344801
step: 290, loss: 0.05140280723571777
step: 300, loss: 0.0785491093993187
step: 310, loss: 0.046047601848840714
step: 320, loss: 0.15891729295253754
step: 330, loss: 0.005801382474601269
step: 340, loss: 0.01797856017947197
step: 350, loss: 0.0019962394144386053
step: 360, loss: 0.002237877808511257
step: 370, loss: 0.0037727304734289646
step: 380, loss: 0.00364652369171381
step: 390, loss: 0.003518258221447468
step: 400, loss: 0.021674279123544693
step: 410, loss: 0.0342731848359108
step: 420, loss: 0.18357323110103607
step: 430, loss: 0.02028162032365799
step: 440, loss: 0.13803927600383759
step: 450, loss: 0.212399423122406
step: 460, loss: 0.09104660153388977
step: 470, loss: 0.03889746591448784
step: 480, loss: 0.006817690096795559
step: 490, loss: 0.0332193560898304
step: 500, loss: 0.009161967784166336
step: 510, loss: 0.0024661861825734377
step: 520, loss: 0.014234182424843311

```

step: 530, loss: 0.015023387968540192
/usr/local/lib/python3.7/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=0.9344337459880789, f1=0.9295380307979467,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02543627843260765
step: 10, loss: 0.007895926013588905
step: 20, loss: 0.004163545090705156
step: 30, loss: 0.004550545942038298
step: 40, loss: 0.001282573095522821
step: 50, loss: 0.005025175400078297
step: 60, loss: 0.09254316240549088
step: 70, loss: 0.002112705260515213
step: 80, loss: 0.025633448734879494
step: 90, loss: 0.04866908863186836
step: 100, loss: 0.010381752625107765
step: 110, loss: 0.024721913039684296
step: 120, loss: 0.04300244152545929
step: 130, loss: 0.014074713923037052
step: 140, loss: 0.002798115834593773
step: 150, loss: 0.0009436198743060231
step: 160, loss: 0.06615009158849716
step: 170, loss: 0.004571512341499329
step: 180, loss: 0.06582914292812347
step: 190, loss: 0.22292016446590424
step: 200, loss: 0.02437349036335945
step: 210, loss: 0.05212430655956268
step: 220, loss: 0.015516535378992558
step: 230, loss: 0.05192519351840019
step: 240, loss: 0.0033310800790786743
step: 250, loss: 0.010705908760428429
step: 260, loss: 0.023934733122587204
step: 270, loss: 0.014135511592030525
step: 280, loss: 0.0022637604270130396
step: 290, loss: 0.004628363065421581
step: 300, loss: 0.003681726288050413
step: 310, loss: 0.19951778650283813
step: 320, loss: 0.0018719421932473779
step: 330, loss: 0.008933614008128643
step: 340, loss: 0.004546693526208401
step: 350, loss: 0.00489721680060029
step: 360, loss: 0.005437348037958145
step: 370, loss: 0.019695919007062912
step: 380, loss: 0.1031794548034668

```

```

step: 390, loss: 0.017162248492240906
step: 400, loss: 0.037092845886945724
step: 410, loss: 0.009044804610311985
step: 420, loss: 0.020571302622556686
step: 430, loss: 0.002165383193641901
step: 440, loss: 0.0008278750465251505
step: 450, loss: 0.2183479368686676
step: 460, loss: 0.007779781706631184
step: 470, loss: 0.0056286342442035675
step: 480, loss: 0.022093497216701508
step: 490, loss: 0.003143079113215208
step: 500, loss: 0.06030775234103203
step: 510, loss: 0.13739104568958282
step: 520, loss: 0.0012227605329826474
step: 530, loss: 0.021909160539507866
/usr/local/lib/python3.7/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=0.9313815187557181, f1=0.925925925925926,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017792920116335154
step: 10, loss: 0.002102920552715659
step: 20, loss: 0.0041181850247085094
step: 30, loss: 0.03450154513120651
step: 40, loss: 0.0024164023343473673
step: 50, loss: 0.02935618907213211
step: 60, loss: 0.0016345734475180507
step: 70, loss: 0.0004341187886893749
step: 80, loss: 0.008161685429513454
step: 90, loss: 0.004945521242916584
step: 100, loss: 0.0011239363811910152
step: 110, loss: 0.0007200157851912081
step: 120, loss: 0.1079743281006813
step: 130, loss: 0.00048131096991710365
step: 140, loss: 0.009125837124884129
step: 150, loss: 0.02200760506093502
step: 160, loss: 0.0009193685255013406
step: 170, loss: 0.04170534014701843
step: 180, loss: 0.040061235427856445
step: 190, loss: 0.007478517480194569
step: 200, loss: 0.0011200743028894067
step: 210, loss: 0.00022609072038903832
step: 220, loss: 0.020101098343729973
step: 230, loss: 0.0007055198075249791
step: 240, loss: 0.05177328735589981

```

```

step: 250, loss: 0.07281693071126938
step: 260, loss: 0.014269497245550156
step: 270, loss: 0.0022992955055087805
step: 280, loss: 0.006246650591492653
step: 290, loss: 0.0025513439904898405
step: 300, loss: 0.00041740352753549814
step: 310, loss: 0.0012088377261534333
step: 320, loss: 0.04237521067261696
step: 330, loss: 0.0013671584893018007
step: 340, loss: 0.00017359759658575058
step: 350, loss: 0.0021774915512651205
step: 360, loss: 0.000228030068683438
step: 370, loss: 0.0195525661110878
step: 380, loss: 0.07582211494445801
step: 390, loss: 0.02748693898320198
step: 400, loss: 0.04000873118638992
step: 410, loss: 0.0022018318995833397
step: 420, loss: 0.005486351437866688
step: 430, loss: 0.012260430492460728
step: 440, loss: 0.021848587319254875
step: 450, loss: 0.027237260714173317
step: 460, loss: 0.002495324471965432
step: 470, loss: 0.16546055674552917
step: 480, loss: 0.007945035584270954
step: 490, loss: 0.0029760103207081556
step: 500, loss: 0.0175702515989542
step: 510, loss: 0.004830487072467804
step: 520, loss: 0.00028079995536245406
step: 530, loss: 0.005022471770644188
/usr/local/lib/python3.7/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=0.9357374017568193, f1=0.9319129226493746,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032543351408094168
step: 10, loss: 0.0007758802967146039
step: 20, loss: 0.0006304754060693085
step: 30, loss: 0.021944014355540276
step: 40, loss: 0.0052810730412602425
step: 50, loss: 0.0009342639823444188
step: 60, loss: 0.01532079093158245
step: 70, loss: 0.028955040499567986
step: 80, loss: 0.005348785314708948
step: 90, loss: 0.01316872425377369
step: 100, loss: 0.011087772436439991

```

```
step: 110, loss: 0.0006838872795924544
step: 120, loss: 0.0008560656569898129
step: 130, loss: 0.002935072174295783
step: 140, loss: 0.10654205083847046
step: 150, loss: 0.0017998652765527368
step: 160, loss: 0.022664586082100868
step: 170, loss: 0.10904897749423981
step: 180, loss: 0.003947189077734947
step: 190, loss: 0.12775427103042603
step: 200, loss: 0.015153573825955391
step: 210, loss: 0.10159645229578018
step: 220, loss: 0.007313641719520092
step: 230, loss: 0.06525719165802002
step: 240, loss: 0.2104329615831375
step: 250, loss: 0.0054368507117033005
step: 260, loss: 0.0012861313298344612
step: 270, loss: 0.009582238271832466
step: 280, loss: 0.04163568094372749
step: 290, loss: 0.006659362930804491
step: 300, loss: 0.002084288513287902
step: 310, loss: 0.0018715556943789124
step: 320, loss: 0.011740499176084995
step: 330, loss: 0.0013578010257333517
step: 340, loss: 0.01367335394024849
step: 350, loss: 0.0006839080015197396
step: 360, loss: 0.010455263778567314
step: 370, loss: 0.24377687275409698
step: 380, loss: 0.0016606093849986792
step: 390, loss: 0.08632919192314148
step: 400, loss: 0.005730229429900646
step: 410, loss: 0.005184752866625786
step: 420, loss: 0.004878136795014143
step: 430, loss: 0.01247997023165226
step: 440, loss: 0.08612404018640518
step: 450, loss: 0.004977701231837273
step: 460, loss: 0.003931726794689894
step: 470, loss: 0.07593853026628494
step: 480, loss: 0.03616431728005409
step: 490, loss: 0.0016352023230865598
step: 500, loss: 0.11569756269454956
step: 510, loss: 0.027433140203356743
step: 520, loss: 0.0053259991109371185
step: 530, loss: 0.003593201283365488
/usr/local/lib/python3.7/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=0.9330232558139535, f1=0.9272641952135147,
best_f1=0.9288664525011474
```


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

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

step: 0, loss: 0.0009230218711309135
step: 10, loss: 0.006502000615000725
step: 20, loss: 0.00017660853336565197
step: 30, loss: 0.10369863361120224
step: 40, loss: 0.0036889705806970596
step: 50, loss: 0.002247895346954465
step: 60, loss: 0.0007602301775477827
step: 70, loss: 0.007409462705254555
step: 80, loss: 0.005737386178225279
step: 90, loss: 0.007122059818357229
step: 100, loss: 0.05324344336986542
step: 110, loss: 0.007488677278161049
step: 120, loss: 0.0038652485236525536
step: 130, loss: 0.0017510614125058055
step: 140, loss: 0.023816946893930435
step: 150, loss: 0.007216922007501125
step: 160, loss: 0.006600687745958567
step: 170, loss: 0.0013261234853416681
step: 180, loss: 0.01630353182554245
step: 190, loss: 0.000561145949177444
step: 200, loss: 0.0010969401337206364
step: 210, loss: 0.010053911246359348
step: 220, loss: 0.0008229310042224824
step: 230, loss: 0.0026157028041779995
step: 240, loss: 0.0006076306453906
step: 250, loss: 0.004408820997923613
step: 260, loss: 0.002611488802358508
step: 270, loss: 0.0009119197493419051
step: 280, loss: 0.0014557165559381247
step: 290, loss: 0.0009737854707054794
step: 300, loss: 0.014442305080592632
step: 310, loss: 0.008031229488551617
step: 320, loss: 0.0037156229373067617
step: 330, loss: 0.004428910091519356
step: 340, loss: 0.046731311827898026
step: 350, loss: 0.20491397380828857
step: 360, loss: 0.043532803654670715
step: 370, loss: 0.0002763580996543169
step: 380, loss: 0.003683590330183506
step: 390, loss: 0.0008981875726021826
step: 400, loss: 0.007332648616284132
step: 410, loss: 0.003608172293752432
step: 420, loss: 0.0015710457228124142
step: 430, loss: 0.0029462631791830063
step: 440, loss: 0.0005870398017577827

```

step: 450, loss: 0.0011707281228154898
step: 460, loss: 0.012724538333714008
step: 470, loss: 0.0008610282675363123
step: 480, loss: 0.009001852944493294
step: 490, loss: 0.0014163326704874635
step: 500, loss: 0.01833421178162098
step: 510, loss: 0.008889159187674522
step: 520, loss: 0.005718624219298363
step: 530, loss: 0.016195738688111305
/usr/local/lib/python3.7/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=0.9360709286047598, f1=0.9313264346190029,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036183572374284267
step: 10, loss: 0.0013667060993611813
step: 20, loss: 0.008106927387416363
step: 30, loss: 0.005471321288496256
step: 40, loss: 0.03908894211053848
step: 50, loss: 0.0016764140455052257
step: 60, loss: 0.00039488085894845426
step: 70, loss: 0.000365017622243613
step: 80, loss: 0.0014208696084097028
step: 90, loss: 0.00202162261120975
step: 100, loss: 0.002209308324381709
step: 110, loss: 0.006611153017729521
step: 120, loss: 0.004054723307490349
step: 130, loss: 0.009013688191771507
step: 140, loss: 0.004395270720124245
step: 150, loss: 0.0022092140279710293
step: 160, loss: 0.03088958188891411
step: 170, loss: 0.0013146671699360013
step: 180, loss: 0.007726968266069889
step: 190, loss: 0.0003690905577968806
step: 200, loss: 0.0011341985082253814
step: 210, loss: 0.015698427334427834
step: 220, loss: 0.1047629565000534
step: 230, loss: 0.0013831929536536336
step: 240, loss: 0.0026448089629411697
step: 250, loss: 0.059804484248161316
step: 260, loss: 0.021290669217705727
step: 270, loss: 0.04755224287509918
step: 280, loss: 0.001276828465051949
step: 290, loss: 0.004321693908423185
step: 300, loss: 0.005605428013950586

```

```

step: 310, loss: 0.013560120016336441
step: 320, loss: 0.0035357968881726265
step: 330, loss: 0.046418629586696625
step: 340, loss: 0.00015353340131696314
step: 350, loss: 0.09236089885234833
step: 360, loss: 0.0013610331807285547
step: 370, loss: 9.03607506188564e-05
step: 380, loss: 0.003365136682987213
step: 390, loss: 0.001176066230982542
step: 400, loss: 0.0007000644109211862
step: 410, loss: 0.00031687552109360695
step: 420, loss: 0.0028277456294745207
step: 430, loss: 0.0022288423497229815
step: 440, loss: 0.0005956141976639628
step: 450, loss: 0.004611152224242687
step: 460, loss: 0.002675495343282819
step: 470, loss: 0.004068596754223108
step: 480, loss: 0.002137131290510297
step: 490, loss: 0.0007766523631289601
step: 500, loss: 0.00424261623993516
step: 510, loss: 0.001428383169695735
step: 520, loss: 0.004592460580170155
step: 530, loss: 0.04049052298069
/usr/local/lib/python3.7/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=0.9346497414198403, f1=0.9262564584311883,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020427550771273673
step: 10, loss: 0.0014476985670626163
step: 20, loss: 0.002810215577483177
step: 30, loss: 0.0008752995636314154
step: 40, loss: 0.0012543199118226767
step: 50, loss: 0.01181510929018259
step: 60, loss: 0.0009829485788941383
step: 70, loss: 0.0007639721734449267
step: 80, loss: 0.0008271416299976408
step: 90, loss: 0.0009156232699751854
step: 100, loss: 0.0010928952833637595
step: 110, loss: 0.004076366778463125
step: 120, loss: 0.00033474076190032065
step: 130, loss: 0.025657907128334045
step: 140, loss: 0.01785808801651001
step: 150, loss: 0.0023428453132510185
step: 160, loss: 0.0003977542801294476

```

```

step: 170, loss: 0.0007523559615947306
step: 180, loss: 0.00032526295399293303
step: 190, loss: 0.0003768807218875736
step: 200, loss: 0.0012011559447273612
step: 210, loss: 0.004116297699511051
step: 220, loss: 0.008744061924517155
step: 230, loss: 0.0004353329713921994
step: 240, loss: 0.005398523062467575
step: 250, loss: 0.0014456266071647406
step: 260, loss: 0.005637768190354109
step: 270, loss: 0.0006657946505583823
step: 280, loss: 0.0007807389483787119
step: 290, loss: 0.0022538688499480486
step: 300, loss: 0.0006679610232822597
step: 310, loss: 0.023074757307767868
step: 320, loss: 0.0004893758450634778
step: 330, loss: 9.616243187338114e-05
step: 340, loss: 0.005563965067267418
step: 350, loss: 0.0007004905492067337
step: 360, loss: 0.025554435327649117
step: 370, loss: 0.00498721981421113
step: 380, loss: 0.0031774400267750025
step: 390, loss: 0.001823842409066856
step: 400, loss: 0.0056471191346645355
step: 410, loss: 0.0009770026663318276
step: 420, loss: 0.012685094960033894
step: 430, loss: 0.0014699576422572136
step: 440, loss: 0.0005353084998205304
step: 450, loss: 0.00879866536706686
step: 460, loss: 0.022784367203712463
step: 470, loss: 0.00048049932229332626
step: 480, loss: 5.97483194724191e-05
step: 490, loss: 0.00701029971241951
step: 500, loss: 0.00017188598576467484
step: 510, loss: 0.0010348324431106448
step: 520, loss: 0.008227905258536339
step: 530, loss: 0.01293687429279089
/usr/local/lib/python3.7/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=0.9335180055401663, f1=0.9291338582677167,
best_f1=0.9288664525011474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003505278844386339
step: 10, loss: 0.00337662803940475
step: 20, loss: 0.01355853769928217

```

step: 30, loss: 0.00012171889102319255
step: 40, loss: 0.013292371295392513
step: 50, loss: 0.0037972277496010065
step: 60, loss: 0.01940314657986164
step: 70, loss: 0.0003634584427345544
step: 80, loss: 0.00014259052113629878
step: 90, loss: 0.00026602353318594396
step: 100, loss: 0.023186199367046356
step: 110, loss: 7.982304668985307e-05
step: 120, loss: 0.0004933247691951692
step: 130, loss: 0.0005827456479892135
step: 140, loss: 0.00020417293126229197
step: 150, loss: 0.005080288741737604
step: 160, loss: 0.0005313549772836268
step: 170, loss: 0.0010804341873154044
step: 180, loss: 0.0036357049830257893
step: 190, loss: 0.001448947936296463
step: 200, loss: 0.001395135186612606
step: 210, loss: 0.00035978714004158974
step: 220, loss: 0.0008323414949700236
step: 230, loss: 0.0007243268773891032
step: 240, loss: 0.0016037641325965524
step: 250, loss: 0.16614004969596863
step: 260, loss: 0.01475249882787466
step: 270, loss: 0.001102876616641879
step: 280, loss: 0.0006275522173382342
step: 290, loss: 0.0035195511300116777
step: 300, loss: 0.002045743865892291
step: 310, loss: 0.01697416976094246
step: 320, loss: 0.03143361583352089
step: 330, loss: 0.0023141277488321066
step: 340, loss: 0.0008734888979233801
step: 350, loss: 0.00029551313491538167
step: 360, loss: 0.0035744013730436563
step: 370, loss: 0.0019134856993332505
step: 380, loss: 0.007881473749876022
step: 390, loss: 0.008702069520950317
step: 400, loss: 0.00020491672330535948
step: 410, loss: 0.02102666161954403
step: 420, loss: 0.0011943705612793565
step: 430, loss: 0.04131095111370087
step: 440, loss: 0.0033626649528741837
step: 450, loss: 0.09794166684150696
step: 460, loss: 0.0019933811854571104
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 470, loss: 0.0036225372459739447
step: 480, loss: 0.025773225352168083
step: 490, loss: 0.001356636406853795

```

step: 500, loss: 0.020604213699698448
step: 510, loss: 0.018803419545292854
step: 520, loss: 0.008267750963568687
step: 530, loss: 0.006365531589835882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9373246024321795, f1=0.9315838800374883,
best_f1=0.9315838800374883
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018377596279606223
step: 10, loss: 0.0004883133224211633
step: 20, loss: 0.0026669041253626347
step: 30, loss: 7.245273445732892e-05
step: 40, loss: 0.0035212193615734577
step: 50, loss: 0.013475008308887482
step: 60, loss: 0.014412999153137207
step: 70, loss: 0.005062179174274206
step: 80, loss: 0.008025445975363255
step: 90, loss: 0.0003232705930713564
step: 100, loss: 0.007601145189255476
step: 110, loss: 0.0026336577720940113
step: 120, loss: 0.0028938171453773975
step: 130, loss: 0.006478547118604183
step: 140, loss: 0.001160553190857172
step: 150, loss: 0.0016225532162934542
step: 160, loss: 0.025890706107020378
step: 170, loss: 0.0007352612446993589
step: 180, loss: 0.0013397295260801911
step: 190, loss: 0.0005043904529884458
step: 200, loss: 0.00011341738718328997
step: 210, loss: 0.00595637783408165
step: 220, loss: 0.036426760256290436
step: 230, loss: 0.010070830583572388
step: 240, loss: 0.0008686622022651136
step: 250, loss: 0.0005072771455161273
step: 260, loss: 0.0009473328245803714
step: 270, loss: 0.17270348966121674
step: 280, loss: 0.003807180095463991
step: 290, loss: 0.011776904575526714
step: 300, loss: 9.994326683226973e-05
step: 310, loss: 0.00107255217153579

```

```

step: 320, loss: 0.002581903710961342
step: 330, loss: 0.0009580028126947582
step: 340, loss: 0.01146735344082117
step: 350, loss: 0.00012344360584393144
step: 360, loss: 0.12640663981437683
step: 370, loss: 0.010212793946266174
step: 380, loss: 0.00045933181536383927
step: 390, loss: 0.0005255842115730047
step: 400, loss: 0.0003264894476160407
step: 410, loss: 0.00017671272507868707
step: 420, loss: 0.0010116543853655457
step: 430, loss: 0.00039567824569530785
step: 440, loss: 0.0006040507578290999
step: 450, loss: 0.03238028660416603
step: 460, loss: 0.0005062703276053071
step: 470, loss: 0.0013256451347842813
step: 480, loss: 0.0005113255465403199
step: 490, loss: 0.0005785905523225665
step: 500, loss: 0.007307484280318022
step: 510, loss: 0.003853969508782029
step: 520, loss: 0.0016352537786588073
step: 530, loss: 0.00016071686695795506
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9375854214123007, f1=0.9285714285714285,
best_f1=0.9285714285714285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01781112514436245
step: 10, loss: 0.0014026411809027195
step: 20, loss: 3.2068790460471064e-05
step: 30, loss: 0.00027536790003068745
step: 40, loss: 0.0004596763465087861
step: 50, loss: 0.09773945063352585
step: 60, loss: 0.00042737636249512434
step: 70, loss: 0.014949873089790344
step: 80, loss: 4.925778557662852e-05
step: 90, loss: 8.085162698989734e-05
step: 100, loss: 0.00015752350736875087
step: 110, loss: 0.00026980185066349804
step: 120, loss: 0.00200975826010108
step: 130, loss: 7.904045924078673e-05

```

```

step: 140, loss: 0.0008983080624602735
step: 150, loss: 0.0008312642457894981
step: 160, loss: 0.0005077594541944563
step: 170, loss: 0.0009875206742435694
step: 180, loss: 5.76547572563868e-05
step: 190, loss: 0.0011490782490000129
step: 200, loss: 0.036937277764081955
step: 210, loss: 0.013656212948262691
step: 220, loss: 0.00022465481015387923
step: 230, loss: 0.00023242815223056823
step: 240, loss: 0.0004887585528194904
step: 250, loss: 0.0016569230938330293
step: 260, loss: 0.006768590770661831
step: 270, loss: 0.0006787653546780348
step: 280, loss: 0.00020997844694647938
step: 290, loss: 0.0005163763416931033
step: 300, loss: 2.8969747290830128e-05
step: 310, loss: 7.686371827730909e-05
step: 320, loss: 3.459786603343673e-05
step: 330, loss: 0.0007596353534609079
step: 340, loss: 0.0002955097588710487
step: 350, loss: 0.0003149946278426796
step: 360, loss: 0.06680188328027725
step: 370, loss: 0.013365935534238815
step: 380, loss: 0.002355885459110141
step: 390, loss: 0.00024144267081283033
step: 400, loss: 0.008928489871323109
step: 410, loss: 0.00131730898283422
step: 420, loss: 8.201503806049004e-05
step: 430, loss: 0.0008981192950159311
step: 440, loss: 0.0011177656706422567
step: 450, loss: 0.0008324099471792579
step: 460, loss: 0.0029804883524775505
step: 470, loss: 2.662056613189634e-05
step: 480, loss: 6.312497134786099e-05
step: 490, loss: 0.0014502127887681127
step: 500, loss: 0.02719845436513424
step: 510, loss: 0.07118108123540878
step: 520, loss: 0.00021438916155602783
step: 530, loss: 0.0008017323561944067
/usr/local/lib/python3.7/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=0.9345622119815669, f1=0.9295644114921224,
best_f1=0.9285714285714285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```


step: 0, loss: 0.00024465747992508113
step: 10, loss: 0.0010972125455737114
step: 20, loss: 0.0008176579722203314
step: 30, loss: 0.004318741615861654
step: 40, loss: 9.364553261548281e-05
step: 50, loss: 0.0009738208027556539
step: 60, loss: 0.00033560709562152624
step: 70, loss: 0.00010108618153026327
step: 80, loss: 0.0005230996757745743
step: 90, loss: 0.0047906544059515
step: 100, loss: 0.00010667349124560133
step: 110, loss: 0.01586259715259075
step: 120, loss: 0.000494497362524271
step: 130, loss: 0.004473199602216482
step: 140, loss: 2.445058271405287e-05
step: 150, loss: 0.00011385582911316305
step: 160, loss: 0.0013846845831722021
step: 170, loss: 4.976492709829472e-05
step: 180, loss: 0.009037848562002182
step: 190, loss: 0.0010237941751256585
step: 200, loss: 0.0021266506519168615
step: 210, loss: 0.00045238915481604636
step: 220, loss: 5.2556559239747e-05
step: 230, loss: 0.00016932415019255131
step: 240, loss: 5.7005683629540727e-05
step: 250, loss: 0.0001566333812661469
step: 260, loss: 4.225858356221579e-05
step: 270, loss: 0.0007947084377519786
step: 280, loss: 7.10362583049573e-05
step: 290, loss: 0.000729748047888279
step: 300, loss: 0.0001847274979809299
step: 310, loss: 0.00034712222986854613
step: 320, loss: 9.481490997131914e-05
step: 330, loss: 0.009078329429030418
step: 340, loss: 0.00025283044669777155
step: 350, loss: 0.0032281142193824053
step: 360, loss: 0.001468396163545549
step: 370, loss: 0.00037576176691800356
step: 380, loss: 0.00037407176569104195
step: 390, loss: 0.002070062793791294
step: 400, loss: 0.01147549506276846
step: 410, loss: 0.007691649720072746
step: 420, loss: 0.0003641662187874317
step: 430, loss: 3.476767233223654e-05
step: 440, loss: 0.0017180141294375062
step: 450, loss: 0.0008193021058104932
step: 460, loss: 0.001909807906486094
step: 470, loss: 0.00011875862401211634

```

step: 480, loss: 0.002669180743396282
step: 490, loss: 0.00012022402370348573
step: 500, loss: 0.008801541291177273
step: 510, loss: 7.965114491526037e-05
step: 520, loss: 1.0128843314305414e-05
step: 530, loss: 0.00038923360989429057
/usr/local/lib/python3.7/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=0.9361900326036331, f1=0.9296037296037295,
best_f1=0.9285714285714285
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:25, 224.14it/s]
load_f1 = 0.9352319706017456
real_f1 = 0.936953520478601
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 195.59it/s]

```

2 BASELINE

2.1 BASELINE STRUCTURED

2.1.1 Beer - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

```

After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.441062867641449
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=0.2828282828282828, f1=0.2947368421052632,
best_f1=0.2947368421052632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44913971424102783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.34146341463414637, f1=0.32352941176470584,
best_f1=0.32352941176470584

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.40065884590148926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3529411764705882, f1=0.4615384615384615,
best_f1=0.4615384615384615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24683749675750732
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3636363636363636, f1=0.3728813559322034,
best_f1=0.3728813559322034
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.27311110496520996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4242424242424242, f1=0.4444444444444445,
best_f1=0.4444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21405752003192902
/usr/local/lib/python3.7/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=0.3010752688172043, f1=0.29213483146067415,
best_f1=0.4444444444444445

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5001158118247986
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.3835616438356164, f1=0.4705882352941177,
best_f1=0.4444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44739460945129395
/usr/local/lib/python3.7/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=0.3943661971830986, f1=0.4705882352941177,
best_f1=0.4444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2879963517189026
/usr/local/lib/python3.7/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=0.39999999999999997, f1=0.4761904761904762,
best_f1=0.4444444444444445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3988233208656311
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5365853658536585, f1=0.5454545454545454,
best_f1=0.5454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4699839949607849
/usr/local/lib/python3.7/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=0.4571428571428571, f1=0.6206896551724138,
best_f1=0.5454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2629753649234772
/usr/local/lib/python3.7/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=0.5238095238095237, f1=0.6285714285714286,
best_f1=0.5454545454545454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23949657380580902
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5555555555555556, f1=0.6451612903225806,
best_f1=0.6451612903225806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1767512708902359
/usr/local/lib/python3.7/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=0.5555555555555556, f1=0.6451612903225806,
best_f1=0.6451612903225806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30019500851631165
/usr/local/lib/python3.7/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=0.5555555555555556, f1=0.6451612903225806,
best_f1=0.6451612903225806
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, 122177.23it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.5106382978723404
real_f1 = 0.46428571428571436
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:31, 139.42it/s]

```

2.1.2 DBLP-ACM - Running the matcher

```

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

```



```

--fp16 \
--save_model

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.5938526391983032

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

step: 10, loss: 0.48562103509902954

step: 20, loss: 0.6333747506141663

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

```

step: 30, loss: 0.3459849953651428
step: 40, loss: 0.3323812186717987
step: 50, loss: 0.5452267527580261
step: 60, loss: 0.47154197096824646
step: 70, loss: 0.43257904052734375
step: 80, loss: 0.5814096927642822
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.07252690941095352
step: 100, loss: 0.33503180742263794
step: 110, loss: 0.3852945864200592
step: 120, loss: 0.1114940196275711
step: 130, loss: 0.009514975361526012
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.27434593439102173
step: 150, loss: 0.16252738237380981
step: 160, loss: 0.16562223434448242
step: 170, loss: 0.4892139136791229
step: 180, loss: 0.3461964726448059
step: 190, loss: 0.04931934177875519
step: 200, loss: 0.09193266928195953
step: 210, loss: 0.020676637068390846
step: 220, loss: 0.025267530232667923
step: 230, loss: 0.008619076572358608
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9647577092511014, f1=0.9678135405105438,
best_f1=0.9678135405105438
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07440400123596191
step: 10, loss: 0.06152735650539398
step: 20, loss: 0.04666152969002724
step: 30, loss: 0.009391325525939465
step: 40, loss: 0.05100557208061218
step: 50, loss: 0.010302886366844177
step: 60, loss: 0.010826312005519867
step: 70, loss: 0.04302038997411728
step: 80, loss: 0.0031378427520394325
step: 90, loss: 0.08385016024112701
step: 100, loss: 0.0058509656228125095
step: 110, loss: 0.13018319010734558
step: 120, loss: 0.00585900666192174

```

```

step: 130, loss: 0.08506608754396439
step: 140, loss: 0.004952284973114729
step: 150, loss: 0.23084203898906708
step: 160, loss: 0.08085712790489197
step: 170, loss: 0.11089438199996948
step: 180, loss: 0.025805726647377014
step: 190, loss: 0.20386505126953125
step: 200, loss: 0.02985413558781147
step: 210, loss: 0.017250224947929382
step: 220, loss: 0.0017987008905038238
step: 230, loss: 0.0017120074480772018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9707865168539327, f1=0.9694224235560589,
best_f1=0.9694224235560589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04583348333835602
step: 10, loss: 0.02133730612695217
step: 20, loss: 0.04529983550310135
step: 30, loss: 0.006853187922388315
step: 40, loss: 0.006692719645798206
step: 50, loss: 0.036768339574337006
step: 60, loss: 0.003625725395977497
step: 70, loss: 0.02437812089920044
step: 80, loss: 0.0005373246967792511
step: 90, loss: 0.0416354201734066
step: 100, loss: 0.013044782914221287
step: 110, loss: 0.0052366554737091064
step: 120, loss: 0.0016216435469686985
step: 130, loss: 0.006635958794504404
step: 140, loss: 0.025638090446591377
step: 150, loss: 0.14262984693050385
step: 160, loss: 0.005493388511240482
step: 170, loss: 0.012542244046926498
step: 180, loss: 0.0406901054084301
step: 190, loss: 0.009869452565908432
step: 200, loss: 0.04750300943851471
step: 210, loss: 0.02468492090702057
step: 220, loss: 0.007924798876047134
step: 230, loss: 0.011605417355895042
/usr/local/lib/python3.7/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=0.9655172413793103, f1=0.9796380090497738,
best_f1=0.9694224235560589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01441898476332426
step: 10, loss: 0.011146421544253826
step: 20, loss: 0.006373291369527578
step: 30, loss: 0.00148735661059618
step: 40, loss: 0.1557079702615738
step: 50, loss: 0.04148606210947037
step: 60, loss: 0.0030208416283130646
step: 70, loss: 0.013582823798060417
step: 80, loss: 0.001978776417672634
step: 90, loss: 0.025033632293343544
step: 100, loss: 0.002396622207015753
step: 110, loss: 0.0037354191299527884
step: 120, loss: 0.003983047790825367
step: 130, loss: 0.03493601083755493
step: 140, loss: 0.0008264645584858954
step: 150, loss: 0.0005122970324009657
step: 160, loss: 0.0009274812182411551
step: 170, loss: 0.011129608377814293
step: 180, loss: 0.010648279450833797
step: 190, loss: 0.0016634612111374736
step: 200, loss: 0.024359749630093575
step: 210, loss: 0.00790405459702015
step: 220, loss: 0.08957234025001526
step: 230, loss: 0.005463816691190004
/usr/local/lib/python3.7/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=0.967525195968645, f1=0.9752808988764046,
best_f1=0.9694224235560589
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10688533633947372
step: 10, loss: 0.011288609355688095
step: 20, loss: 0.040032584220170975
step: 30, loss: 0.11199728399515152
step: 40, loss: 0.012127633206546307
step: 50, loss: 0.05533739551901817
step: 60, loss: 0.017913270741701126
step: 70, loss: 0.005926681216806173
step: 80, loss: 0.011351799592375755

```

```

step: 90, loss: 0.03774731978774071
step: 100, loss: 0.0004126751737203449
step: 110, loss: 0.05606576055288315
step: 120, loss: 0.046321019530296326
step: 130, loss: 0.00117921968922019
step: 140, loss: 0.0007618131348863244
step: 150, loss: 0.007205943111330271
step: 160, loss: 0.06658142060041428
step: 170, loss: 0.017257103696465492
step: 180, loss: 0.01527467742562294
step: 190, loss: 0.1411631852388382
step: 200, loss: 0.007597208023071289
step: 210, loss: 0.0028288105968385935
step: 220, loss: 0.0012273446191102266
step: 230, loss: 0.01653279922902584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774266365688488, f1=0.9808773903262092,
best_f1=0.9808773903262092
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006522138719446957
step: 10, loss: 0.0016581331146880984
step: 20, loss: 0.0032140163239091635
step: 30, loss: 0.0010925107635557652
step: 40, loss: 0.0017818816704675555
step: 50, loss: 0.0024088609497994184
step: 60, loss: 0.005603129975497723
step: 70, loss: 0.004805885720998049
step: 80, loss: 0.004340043291449547
step: 90, loss: 0.018218912184238434
step: 100, loss: 0.0055126831866800785
step: 110, loss: 0.002260128501802683
step: 120, loss: 0.0003397281689103693
step: 130, loss: 0.010390058159828186
step: 140, loss: 0.004406701307743788
step: 150, loss: 0.008397446013987064
step: 160, loss: 0.001289120758883655
step: 170, loss: 0.00026617993717081845
step: 180, loss: 0.00014477618969976902
step: 190, loss: 0.0002866294817067683
step: 200, loss: 0.0057721324265003204

```

```

step: 210, loss: 0.00039975077379494905
step: 220, loss: 0.006931490730494261
step: 230, loss: 0.0006662625819444656
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821428571428571, f1=0.9798657718120806,
best_f1=0.9798657718120806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007846711087040603
step: 10, loss: 0.00040572983562014997
step: 20, loss: 0.00029023404931649566
step: 30, loss: 0.00039074523374438286
step: 40, loss: 0.0006361599662341177
step: 50, loss: 0.0004270822973921895
step: 60, loss: 0.0005432607140392065
step: 70, loss: 0.0004036369500681758
step: 80, loss: 0.0003189498675055802
step: 90, loss: 0.0002844003611244261
step: 100, loss: 0.0003140360931865871
step: 110, loss: 0.00027684870292432606
step: 120, loss: 0.0003468646318651736
step: 130, loss: 0.00022905458172317594
step: 140, loss: 0.00018452052609063685
step: 150, loss: 0.0006285668932832778
step: 160, loss: 0.0003802869759965688
step: 170, loss: 0.011521594598889351
step: 180, loss: 0.0007365313940681517
step: 190, loss: 0.002512479666620493
step: 200, loss: 0.007209653966128826
step: 210, loss: 0.0023914561606943607
step: 220, loss: 0.0011875564232468605
step: 230, loss: 0.05240686610341072
/usr/local/lib/python3.7/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=0.9810055865921787, f1=0.9842342342342343,
best_f1=0.9798657718120806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00752011826261878

```

```

step: 10, loss: 0.0030739977955818176
step: 20, loss: 0.0009172834688797593
step: 30, loss: 0.0005402707611210644
step: 40, loss: 0.0006064899498596787
step: 50, loss: 0.0003634823951870203
step: 60, loss: 0.13788209855556488
step: 70, loss: 0.0014262925833463669
step: 80, loss: 0.04880150035023689
step: 90, loss: 0.0004771807580254972
step: 100, loss: 0.0032821090426295996
step: 110, loss: 0.002207112265750766
step: 120, loss: 0.0014656824059784412
step: 130, loss: 0.0006125461659394205
step: 140, loss: 0.00041506264824420214
step: 150, loss: 0.007384235970675945
step: 160, loss: 0.0008938332321122289
step: 170, loss: 0.043737709522247314
step: 180, loss: 0.0010104079265147448
step: 190, loss: 0.06086602061986923
step: 200, loss: 0.006020913366228342
step: 210, loss: 0.0017920290119946003
step: 220, loss: 0.0009266135166399181
step: 230, loss: 0.0014651952078565955
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.984304932735426,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00043562453356571496
step: 10, loss: 0.00043169857235625386
step: 20, loss: 0.0005011096363887191
step: 30, loss: 0.0052239540964365005
step: 40, loss: 0.00028876078431494534
step: 50, loss: 0.0031968161929398775
step: 60, loss: 0.00028994851163588464
step: 70, loss: 0.003945299424231052
step: 80, loss: 0.00046493910485878587
step: 90, loss: 0.0071648601442575455
step: 100, loss: 0.0002216530847363174
step: 110, loss: 0.002970176748931408
step: 120, loss: 0.0029049876611679792

```

```

step: 130, loss: 0.0005279594333842397
step: 140, loss: 0.0006539154564961791
step: 150, loss: 0.015003843232989311
step: 160, loss: 0.0009324262500740588
step: 170, loss: 0.00028157129418104887
step: 180, loss: 0.0005178634310141206
step: 190, loss: 0.00030248588882386684
step: 200, loss: 0.00032683523022569716
step: 210, loss: 0.003941940143704414
step: 220, loss: 0.0004904085071757436
step: 230, loss: 0.0004538496141321957
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9898305084745763,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003860544820781797
step: 10, loss: 0.0006416024407371879
step: 20, loss: 0.0004023470974061638
step: 30, loss: 0.00019442591292317957
step: 40, loss: 0.005896559450775385
step: 50, loss: 0.00014363403897732496
step: 60, loss: 0.00036162816104479134
step: 70, loss: 0.004156539682298899
step: 80, loss: 0.0005575005779974163
step: 90, loss: 0.0003768513852264732
step: 100, loss: 0.0009303194237872958
step: 110, loss: 0.0025942407082766294
step: 120, loss: 0.00013771136582363397
step: 130, loss: 0.00029373710276558995
step: 140, loss: 0.0001630053884582594
step: 150, loss: 0.0006868263008072972
step: 160, loss: 0.00012848214828409255
step: 170, loss: 0.00021655880846083164
step: 180, loss: 0.0015735370106995106
step: 190, loss: 0.00032815386657603085
step: 200, loss: 0.0004162569239269942
step: 210, loss: 0.002785316202789545
step: 220, loss: 0.0001751650415826589
step: 230, loss: 0.0002907566668000072
/usr/local/lib/python3.7/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=0.9840546697038726, f1=0.9829351535836178,
best_f1=0.984304932735426

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014116255624685436
step: 10, loss: 0.0003339835675433278
step: 20, loss: 0.00020543599384836853
step: 30, loss: 0.0003020930744241923
step: 40, loss: 4.196958980173804e-05
step: 50, loss: 0.0001126006172853522
step: 60, loss: 0.00024417118402197957
step: 70, loss: 0.0002492843195796013
step: 80, loss: 0.0009658181807026267
step: 90, loss: 0.01749914512038231
step: 100, loss: 0.0003958838642574847
step: 110, loss: 0.003406773554161191
step: 120, loss: 0.00015309237642213702
step: 130, loss: 0.00021187915990594774
step: 140, loss: 0.00015639662160538137
step: 150, loss: 0.00019029249961022288
step: 160, loss: 0.20853757858276367
step: 170, loss: 0.019206320866942406
step: 180, loss: 0.001155993901193142
step: 190, loss: 0.0012593016726896167
step: 200, loss: 0.001793174771592021
step: 210, loss: 0.00021524530893657357
step: 220, loss: 0.00017384484817739576
step: 230, loss: 0.0001509020512457937
/usr/local/lib/python3.7/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=0.9818181818181818, f1=0.9864253393665158,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003904259647242725
step: 10, loss: 0.0001932278973981738
step: 20, loss: 0.003988394979387522
step: 30, loss: 0.011612806469202042
step: 40, loss: 0.0004978320794180036
step: 50, loss: 0.0006217342452146113
step: 60, loss: 0.0010496187023818493
step: 70, loss: 0.001280471682548523
step: 80, loss: 0.0022960505448281765
step: 90, loss: 0.006038527935743332
step: 100, loss: 0.00011685074423439801
step: 110, loss: 9.947604849003255e-05
step: 120, loss: 0.0020342052448540926

```

```

step: 130, loss: 0.00028009992092847824
step: 140, loss: 0.0005118577391840518
step: 150, loss: 0.0003010753425769508
step: 160, loss: 0.0035787662491202354
step: 170, loss: 0.0031522370409220457
step: 180, loss: 0.00014434251352213323
step: 190, loss: 0.002723670331761241
step: 200, loss: 0.0010627801530063152
step: 210, loss: 0.001224437728524208
step: 220, loss: 0.0088764987885952
step: 230, loss: 0.0013648371677845716
/usr/local/lib/python3.7/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=0.9797297297297298, f1=0.987598647125141,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007183703128248453
step: 10, loss: 0.00027008456527255476
step: 20, loss: 0.00035223105805926025
step: 30, loss: 5.138126289239153e-05
step: 40, loss: 0.00022314206580631435
step: 50, loss: 0.003908944316208363
step: 60, loss: 0.00010816259600687772
step: 70, loss: 0.00046603637747466564
step: 80, loss: 0.0006485520279966295
step: 90, loss: 0.000328767579048872
step: 100, loss: 0.002021492924541235
step: 110, loss: 0.000874347286298871
step: 120, loss: 0.00021527479111682624
step: 130, loss: 0.00018399101099930704
step: 140, loss: 0.00020993981161154807
step: 150, loss: 0.00023829842393752187
step: 160, loss: 0.020882701501250267
step: 170, loss: 0.00011049041495425627
step: 180, loss: 0.0041602598503232
step: 190, loss: 0.00016600106027908623
step: 200, loss: 6.94645059411414e-05
step: 210, loss: 0.00017572996148373932
step: 220, loss: 0.00014678132720291615
step: 230, loss: 0.00013137958012521267
/usr/local/lib/python3.7/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=0.9797752808988766, f1=0.9820627802690582,
best_f1=0.984304932735426

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.212619286496192e-05
step: 10, loss: 5.703007263946347e-05
step: 20, loss: 0.00010730259964475408
step: 30, loss: 0.00010977705824188888
step: 40, loss: 0.00021776757785119116
step: 50, loss: 8.339035412063822e-05
step: 60, loss: 0.00011840133083751425
step: 70, loss: 0.006662969943135977
step: 80, loss: 7.800775347277522e-05
step: 90, loss: 0.00018947769422084093
step: 100, loss: 8.573298691771924e-05
step: 110, loss: 9.352356573799625e-05
step: 120, loss: 3.54822441295255e-05
step: 130, loss: 0.0006350176990963519
step: 140, loss: 8.940324914874509e-05
step: 150, loss: 4.313972021918744e-05
step: 160, loss: 6.75059127388522e-05
step: 170, loss: 6.542647315654904e-05
step: 180, loss: 8.189469372155145e-05
step: 190, loss: 0.0008324403315782547
step: 200, loss: 0.0003122072957921773
step: 210, loss: 0.0001267645857296884
step: 220, loss: 0.00011879367957590148
step: 230, loss: 5.167825293028727e-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=0.9830890642615557, f1=0.9864559819413092,
best_f1=0.984304932735426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017469908925704658
step: 10, loss: 0.00014470572932623327
step: 20, loss: 0.002603573724627495
step: 30, loss: 0.0001004577788989991
step: 40, loss: 9.141235932474956e-05
step: 50, loss: 0.00016185594722628593
step: 60, loss: 0.032939814031124115
step: 70, loss: 0.00015657702169846743
step: 80, loss: 7.768510840833187e-05
step: 90, loss: 6.618574116146192e-05
step: 100, loss: 4.79082009405829e-05
step: 110, loss: 0.00017061250400729477
step: 120, loss: 0.0010037440806627274

```

```

step: 130, loss: 0.00015031409566290677
step: 140, loss: 0.0024686024989932775
step: 150, loss: 0.0004430816334206611
step: 160, loss: 0.00013885246880818158
step: 170, loss: 3.7993006117176265e-05
step: 180, loss: 8.610275835962966e-05
step: 190, loss: 0.00010386539361206815
step: 200, loss: 0.00016247938037849963
step: 210, loss: 0.006296322215348482
step: 220, loss: 0.0001723758177831769
step: 230, loss: 0.00015021975559648126
/usr/local/lib/python3.7/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=0.9841269841269841, f1=0.9875706214689265,
best_f1=0.984304932735426
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:17, 145.10it/s]
load_f1 = 0.9865771812080537

```

```

real_f1 = 0.9843749999999999
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 136.40it/s]

```

2.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.629680335521698
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.43175941705703735
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.3114365339279175
step: 30, loss: 0.3141498565673828
step: 40, loss: 0.3172353506088257
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.5808732509613037
step: 60, loss: 0.164769247174263
step: 70, loss: 0.17624874413013458
step: 80, loss: 0.2139173150062561
step: 90, loss: 0.12726716697216034
step: 100, loss: 0.33401089906692505
step: 110, loss: 0.14215819537639618
step: 120, loss: 0.15814858675003052
step: 130, loss: 0.29725322127342224
step: 140, loss: 0.17667295038700104
step: 150, loss: 0.12264479696750641
step: 160, loss: 0.16256290674209595
step: 170, loss: 0.11906136572360992
step: 180, loss: 0.1393405795097351
step: 190, loss: 0.13497178256511688
step: 200, loss: 0.10492784529924393
step: 210, loss: 0.04792417585849762
step: 220, loss: 0.0836707353591919
step: 230, loss: 0.20006729662418365
step: 240, loss: 0.06162654608488083
step: 250, loss: 0.09375179558992386
step: 260, loss: 0.4345814287662506
step: 270, loss: 0.27628207206726074

```

```

step: 280, loss: 0.07775543630123138
step: 290, loss: 0.14721234142780304
step: 300, loss: 0.12682917714118958
step: 310, loss: 0.19645623862743378
step: 320, loss: 0.11628822237253189
step: 330, loss: 0.09139484167098999
step: 340, loss: 0.6751250624656677
step: 350, loss: 0.21702203154563904
step: 360, loss: 0.04478643834590912
step: 370, loss: 0.007595019415020943
step: 380, loss: 0.31458520889282227
step: 390, loss: 0.05240287259221077
step: 400, loss: 0.03738105297088623
step: 410, loss: 0.24271635711193085
step: 420, loss: 0.088417649269104
step: 430, loss: 0.010109161026775837
step: 440, loss: 0.06085675582289696
step: 450, loss: 0.13612860441207886
step: 460, loss: 0.027313554659485817
step: 470, loss: 0.07523215562105179
step: 480, loss: 0.10194891691207886
step: 490, loss: 0.12896791100502014
step: 500, loss: 0.04108947142958641
step: 510, loss: 0.027824928984045982
step: 520, loss: 0.11346544325351715
step: 530, loss: 0.09527228027582169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9166666666666666, f1=0.9150507848568791,
best_f1=0.9150507848568791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11982619017362595
step: 10, loss: 0.0931345671415329
step: 20, loss: 0.05803162232041359
step: 30, loss: 0.06155838072299957
step: 40, loss: 0.02871089242398739
step: 50, loss: 0.08321170508861542
step: 60, loss: 0.05103663355112076
step: 70, loss: 0.04026106745004654
step: 80, loss: 0.06575610488653183
step: 90, loss: 0.25028541684150696

```

```
step: 100, loss: 0.17229235172271729
step: 110, loss: 0.052205588668584824
step: 120, loss: 0.09621263295412064
step: 130, loss: 0.013427119702100754
step: 140, loss: 0.09291041642427444
step: 150, loss: 0.10348770022392273
step: 160, loss: 0.07038123905658722
step: 170, loss: 0.036126021295785904
step: 180, loss: 0.025622976943850517
step: 190, loss: 0.05638252571225166
step: 200, loss: 0.10015246272087097
step: 210, loss: 0.02452709525823593
step: 220, loss: 0.010032973252236843
step: 230, loss: 0.28006142377853394
step: 240, loss: 0.08966798335313797
step: 250, loss: 0.021270956844091415
step: 260, loss: 0.08263510465621948
step: 270, loss: 0.010080794803798199
step: 280, loss: 0.09169990569353104
step: 290, loss: 0.08845221996307373
step: 300, loss: 0.04984324425458908
step: 310, loss: 0.12109357863664627
step: 320, loss: 0.2539539635181427
step: 330, loss: 0.023224351927638054
step: 340, loss: 0.19416488707065582
step: 350, loss: 0.026232430711388588
step: 360, loss: 0.04343106970191002
step: 370, loss: 0.015620467253029346
step: 380, loss: 0.2139187902212143
step: 390, loss: 0.010273735038936138
step: 400, loss: 0.1314421445131302
step: 410, loss: 0.0559769906103611
step: 420, loss: 0.08841246366500854
step: 430, loss: 0.21161171793937683
step: 440, loss: 0.011045346967875957
step: 450, loss: 0.0349871963262558
step: 460, loss: 0.024622049182653427
step: 470, loss: 0.18505162000656128
step: 480, loss: 0.027182897552847862
step: 490, loss: 0.054943520575761795
step: 500, loss: 0.003097415203228593
step: 510, loss: 0.06046804413199425
step: 520, loss: 0.4317646622657776
step: 530, loss: 0.11884469538927078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```


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=0.9274563820018366, f1=0.9305492510213346,

best_f1=0.9305492510213346

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

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

step: 0, loss: 0.13948920369148254

step: 10, loss: 0.04884772747755051

step: 20, loss: 0.005324604455381632

step: 30, loss: 0.054531145840883255

step: 40, loss: 0.021364878863096237

step: 50, loss: 0.1250738501548767

step: 60, loss: 0.06266967952251434

step: 70, loss: 0.022289399057626724

step: 80, loss: 0.026159878820180893

step: 90, loss: 0.03504432737827301

step: 100, loss: 0.21271854639053345

step: 110, loss: 0.11461307108402252

step: 120, loss: 0.06803617626428604

step: 130, loss: 0.04191787913441658

step: 140, loss: 0.012765977531671524

step: 150, loss: 0.05366409942507744

step: 160, loss: 0.027216864749789238

step: 170, loss: 0.06726095825433731

step: 180, loss: 0.017621180042624474

step: 190, loss: 0.00556514086201787

step: 200, loss: 0.06145872548222542

step: 210, loss: 0.0548914298415184

step: 220, loss: 0.029523909091949463

step: 230, loss: 0.048098284751176834

step: 240, loss: 0.017675410956144333

step: 250, loss: 0.09265591949224472

step: 260, loss: 0.017129043117165565

step: 270, loss: 0.048690445721149445

step: 280, loss: 0.002819737186655402

step: 290, loss: 0.14290279150009155

step: 300, loss: 0.1783103346824646

step: 310, loss: 0.08049929887056351

step: 320, loss: 0.1799211949110031

step: 330, loss: 0.006874922197312117

step: 340, loss: 0.01734117977321148

step: 350, loss: 0.0344221256673336

step: 360, loss: 0.011758668348193169

step: 370, loss: 0.1692887246608734

step: 380, loss: 0.041850849986076355

step: 390, loss: 0.017205998301506042

```

step: 400, loss: 0.2860925495624542
step: 410, loss: 0.13197818398475647
step: 420, loss: 0.034276243299245834
step: 430, loss: 0.03538311645388603
step: 440, loss: 0.17488279938697815
step: 450, loss: 0.06430917978286743
step: 460, loss: 0.15020082890987396
step: 470, loss: 0.0038180879782885313
step: 480, loss: 0.11265254020690918
step: 490, loss: 0.0072954026982188225
step: 500, loss: 0.055973898619413376
step: 510, loss: 0.026306306943297386
step: 520, loss: 0.0029277147259563208
step: 530, loss: 0.010993215255439281
/usr/local/lib/python3.7/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=0.9254426840633737, f1=0.9227906976744187,
best_f1=0.9305492510213346
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03301657363772392
step: 10, loss: 0.007326907478272915
step: 20, loss: 0.20568591356277466
step: 30, loss: 0.03164760768413544
step: 40, loss: 0.015300288796424866
step: 50, loss: 0.024897266179323196
step: 60, loss: 0.007034568581730127
step: 70, loss: 0.06444208323955536
step: 80, loss: 0.19308751821517944
step: 90, loss: 0.12215981632471085
step: 100, loss: 0.0023580545093864202
step: 110, loss: 0.17240601778030396
step: 120, loss: 0.010591224767267704
step: 130, loss: 0.043879132717847824
step: 140, loss: 0.00943259708583355
step: 150, loss: 0.01214094553142786
step: 160, loss: 0.005643244832754135
step: 170, loss: 0.0072900960221886635
step: 180, loss: 0.11389496177434921
step: 190, loss: 0.1266767978668213
step: 200, loss: 0.023490577936172485
step: 210, loss: 0.0062323459424078465
step: 220, loss: 0.02788694016635418
step: 230, loss: 0.032669197767972946
step: 240, loss: 0.025789543986320496
step: 250, loss: 0.04843505099415779

```

```

step: 260, loss: 0.012255621142685413
step: 270, loss: 0.05958543345332146
step: 280, loss: 0.058479294180870056
step: 290, loss: 0.012399320490658283
step: 300, loss: 0.005856127478182316
step: 310, loss: 0.03304709494113922
step: 320, loss: 0.19704462587833405
step: 330, loss: 0.019253037869930267
step: 340, loss: 0.016709014773368835
step: 350, loss: 0.02400650829076767
step: 360, loss: 0.12310384958982468
step: 370, loss: 0.013060852885246277
step: 380, loss: 0.003944440744817257
step: 390, loss: 0.009186454117298126
step: 400, loss: 0.03461157903075218
step: 410, loss: 0.007338410243391991
step: 420, loss: 0.02195519581437111
step: 430, loss: 0.05311669036746025
step: 440, loss: 0.004142853897064924
step: 450, loss: 0.053425133228302
step: 460, loss: 0.036596182733774185
step: 470, loss: 0.0055281417444348335
step: 480, loss: 0.1919451802968979
step: 490, loss: 0.010945295915007591
step: 500, loss: 0.31365180015563965
step: 510, loss: 0.07643971592187881
step: 520, loss: 0.01100990641862154
step: 530, loss: 0.1736065149307251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9322671683913453, f1=0.9257683215130024,
best_f1=0.9257683215130024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002572464058175683
step: 10, loss: 0.0140226436778903
step: 20, loss: 0.06484180688858032
step: 30, loss: 0.19269324839115143
step: 40, loss: 0.004143074154853821
step: 50, loss: 0.12802422046661377
step: 60, loss: 0.012447649613022804
step: 70, loss: 0.0030825159046798944

```

```
step: 80, loss: 0.002635106211528182
step: 90, loss: 0.002633033087477088
step: 100, loss: 0.10956881195306778
step: 110, loss: 0.02352232113480568
step: 120, loss: 0.1766478419303894
step: 130, loss: 0.014576088637113571
step: 140, loss: 0.005929882638156414
step: 150, loss: 0.008844430558383465
step: 160, loss: 0.008799230679869652
step: 170, loss: 0.01874133199453354
step: 180, loss: 0.0018626826349645853
step: 190, loss: 0.006171423476189375
step: 200, loss: 0.004912331700325012
step: 210, loss: 0.0016239344840869308
step: 220, loss: 0.039661336690187454
step: 230, loss: 0.004136991687119007
step: 240, loss: 0.006119880825281143
step: 250, loss: 0.15711934864521027
step: 260, loss: 0.010936273261904716
step: 270, loss: 0.009327058680355549
step: 280, loss: 0.004093049094080925
step: 290, loss: 0.09156054258346558
step: 300, loss: 0.21820537745952606
step: 310, loss: 0.021249903365969658
step: 320, loss: 0.07006686180830002
step: 330, loss: 0.002402074169367552
step: 340, loss: 0.008761836215853691
step: 350, loss: 0.11817622929811478
step: 360, loss: 0.0016862807096913457
step: 370, loss: 0.0019462648779153824
step: 380, loss: 0.003551474306732416
step: 390, loss: 0.004013270139694214
step: 400, loss: 0.012776031158864498
step: 410, loss: 0.00879279151558876
step: 420, loss: 0.08786216378211975
step: 430, loss: 0.014706517569720745
step: 440, loss: 0.000697198323905468
step: 450, loss: 0.001955314539372921
step: 460, loss: 0.02999928593635559
step: 470, loss: 0.02585727721452713
step: 480, loss: 0.0030410755425691605
step: 490, loss: 0.014846544712781906
step: 500, loss: 0.09733512997627258
step: 510, loss: 0.032935258001089096
step: 520, loss: 0.3106621503829956
step: 530, loss: 0.004235659260302782
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9324817518248174, f1=0.9230769230769231,
best_f1=0.9230769230769231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.027654124423861504
step: 10, loss: 0.001140695414505899
step: 20, loss: 0.014146311208605766
step: 30, loss: 0.008123229257762432
step: 40, loss: 0.06868819147348404
step: 50, loss: 0.0769839808344841
step: 60, loss: 0.03311145678162575
step: 70, loss: 0.03915689140558243
step: 80, loss: 0.004384187050163746
step: 90, loss: 0.0006568501703441143
step: 100, loss: 0.34359270334243774
step: 110, loss: 0.00344346696510911
step: 120, loss: 0.00908522680401802
step: 130, loss: 0.0018734189216047525
step: 140, loss: 0.001078462926670909
step: 150, loss: 0.0013834593119099736
step: 160, loss: 0.08099190890789032
step: 170, loss: 0.00931699387729168
step: 180, loss: 0.0014127466129139066
step: 190, loss: 0.06263530999422073
step: 200, loss: 0.017343275249004364
step: 210, loss: 0.012001263909041882
step: 220, loss: 0.027164852246642113
step: 230, loss: 0.0319710448384285
step: 240, loss: 0.03500097990036011
step: 250, loss: 0.1980077177286148
step: 260, loss: 0.03707345202565193
step: 270, loss: 0.0030181892216205597
step: 280, loss: 0.0042432076297700405
step: 290, loss: 0.001693011145107448
step: 300, loss: 0.001333201420493424
step: 310, loss: 0.0141260651871562
step: 320, loss: 0.0001508362329332158
step: 330, loss: 0.0217172522097826
step: 340, loss: 0.00047770075616426766
step: 350, loss: 0.029022572562098503
step: 360, loss: 0.08943935483694077
step: 370, loss: 0.004000859800726175

```

```

step: 380, loss: 0.005511286202818155
step: 390, loss: 0.01867753639817238
step: 400, loss: 0.009883249178528786
step: 410, loss: 0.005600467789918184
step: 420, loss: 0.01644296385347843
step: 430, loss: 0.00017543771537020802
step: 440, loss: 0.0017064149724319577
step: 450, loss: 0.05015953257679939
step: 460, loss: 0.0715038850903511
step: 470, loss: 0.007554801646620035
step: 480, loss: 0.11457246541976929
step: 490, loss: 0.07369835674762726
step: 500, loss: 0.005708020646125078
step: 510, loss: 0.22516348958015442
step: 520, loss: 0.004711732268333435
step: 530, loss: 0.06755684316158295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9353233830845772, f1=0.9276285844333181,
best_f1=0.9276285844333181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01390145719051361
step: 10, loss: 0.002785701537504792
step: 20, loss: 0.002032115589827299
step: 30, loss: 0.023734163492918015
step: 40, loss: 0.0012060145381838083
step: 50, loss: 0.010164277628064156
step: 60, loss: 0.0039795516058802605
step: 70, loss: 0.0070814467035233974
step: 80, loss: 0.000881667947396636
step: 90, loss: 0.00025944216758944094
step: 100, loss: 0.0003486165660433471
step: 110, loss: 0.0005724957445636392
step: 120, loss: 0.0015747197903692722
step: 130, loss: 0.00021095202828291804
step: 140, loss: 0.0003671112353913486
step: 150, loss: 0.010206182487308979
step: 160, loss: 0.004664258100092411
step: 170, loss: 0.008610139600932598
step: 180, loss: 0.06310689449310303
step: 190, loss: 0.0013889153487980366

```

```

step: 200, loss: 0.0015903160674497485
step: 210, loss: 0.18015450239181519
step: 220, loss: 0.05579695478081703
step: 230, loss: 0.0006810504128225148
step: 240, loss: 0.0008808057173155248
step: 250, loss: 0.0013751272344961762
step: 260, loss: 0.012419505044817924
step: 270, loss: 0.0002894254575949162
step: 280, loss: 0.0010982801904901862
step: 290, loss: 0.003571548964828253
step: 300, loss: 0.0009217915358021855
step: 310, loss: 0.0008437392534688115
step: 320, loss: 0.018662523478269577
step: 330, loss: 0.0042294468730688095
step: 340, loss: 0.0014449672307819128
step: 350, loss: 0.027630720287561417
step: 360, loss: 0.001372880069538951
step: 370, loss: 0.004483519122004509
step: 380, loss: 0.0026510946918278933
step: 390, loss: 0.035990260541439056
step: 400, loss: 0.13416755199432373
step: 410, loss: 0.002907323418185115
step: 420, loss: 0.004823458846658468
step: 430, loss: 0.015850527212023735
step: 440, loss: 0.0007233255309984088
step: 450, loss: 0.0045258477330207825
step: 460, loss: 0.07981184870004654
step: 470, loss: 0.136990487575531
step: 480, loss: 0.0005697853048332036
step: 490, loss: 0.006618824787437916
step: 500, loss: 0.002339256927371025
step: 510, loss: 0.00011387075210222974
step: 520, loss: 0.0060467179864645
step: 530, loss: 0.0008406029082834721
/usr/local/lib/python3.7/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=0.9271217712177123, f1=0.9222222222222223,
best_f1=0.9276285844333181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018198990437667817
step: 10, loss: 0.00027575617423281074
step: 20, loss: 6.642167136305943e-05
step: 30, loss: 0.00031711251358501613
step: 40, loss: 0.0003063795738853514
step: 50, loss: 0.0024500174913555384

```

step: 60, loss: 0.00026073423214256763
step: 70, loss: 0.0032186219468712807
step: 80, loss: 0.01452922448515892
step: 90, loss: 0.0003915579291060567
step: 100, loss: 0.002798357978463173
step: 110, loss: 0.0012633127626031637
step: 120, loss: 0.0058888322673738
step: 130, loss: 0.0007888861000537872
step: 140, loss: 0.012801758013665676
step: 150, loss: 0.00038041206425987184
step: 160, loss: 0.11782947182655334
step: 170, loss: 0.019421564415097237
step: 180, loss: 0.014479480683803558
step: 190, loss: 0.008545000106096268
step: 200, loss: 0.002792560262605548
step: 210, loss: 0.03430136665701866
step: 220, loss: 0.011023357510566711
step: 230, loss: 0.14374759793281555
step: 240, loss: 0.04527280852198601
step: 250, loss: 0.0008525883313268423
step: 260, loss: 0.0002246765943709761
step: 270, loss: 0.05842183530330658
step: 280, loss: 0.00018192263087257743
step: 290, loss: 0.0012964651687070727
step: 300, loss: 8.91543022589758e-05
step: 310, loss: 0.0002812313032336533
step: 320, loss: 0.0005671947146765888
step: 330, loss: 0.0003124113427475095
step: 340, loss: 0.017799807712435722
step: 350, loss: 0.006486350204795599
step: 360, loss: 0.013518177904188633
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 370, loss: 0.13877475261688232
step: 380, loss: 0.0009718420915305614
step: 390, loss: 0.013689447194337845
step: 400, loss: 0.0003272646863479167
step: 410, loss: 0.0017785237869247794
step: 420, loss: 0.0002731347631197423
step: 430, loss: 0.0002774264430627227
step: 440, loss: 0.0015118424780666828
step: 450, loss: 0.001155554666183889
step: 460, loss: 0.007251799572259188
step: 470, loss: 0.05374271795153618
step: 480, loss: 0.003367780474945903
step: 490, loss: 0.0019903325010091066
step: 500, loss: 0.0017816643230617046
step: 510, loss: 0.12376514077186584
step: 520, loss: 0.0002901339903473854


```
step: 530, loss: 0.003823414910584688
/usr/local/lib/python3.7/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=0.9293504030346136, f1=0.9246901811248809,
best_f1=0.9276285844333181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002054545795544982
step: 10, loss: 0.0053407493978738785
step: 20, loss: 0.0005148958298377693
step: 30, loss: 0.0856022909283638
step: 40, loss: 0.00042664946522563696
step: 50, loss: 0.015413272194564342
step: 60, loss: 0.000474046275485307
step: 70, loss: 0.0003770636103581637
step: 80, loss: 0.03222255781292915
step: 90, loss: 0.03281306475400925
step: 100, loss: 0.00027435977244749665
step: 110, loss: 0.005108838900923729
step: 120, loss: 0.0004660889389924705
step: 130, loss: 0.0010150981834158301
step: 140, loss: 0.0003449978830758482
step: 150, loss: 0.00017609127098694444
step: 160, loss: 0.26260069012641907
step: 170, loss: 0.010313811711966991
step: 180, loss: 0.01179942861199379
step: 190, loss: 0.0009462814196012914
step: 200, loss: 0.01601625233888626
step: 210, loss: 0.1058863177895546
step: 220, loss: 0.007108369842171669
step: 230, loss: 0.0019244066206738353
step: 240, loss: 0.0008907892042770982
step: 250, loss: 0.0026847142726182938
step: 260, loss: 0.08940631151199341
step: 270, loss: 0.001300258096307516
step: 280, loss: 0.0027262119110673666
step: 290, loss: 0.0003399818961042911
step: 300, loss: 0.028862399980425835
step: 310, loss: 0.041045498102903366
step: 320, loss: 0.0044095744378864765
step: 330, loss: 0.029832903295755386
step: 340, loss: 0.0033136142883449793
step: 350, loss: 0.010368488729000092
step: 360, loss: 0.0003796473320107907
step: 370, loss: 0.017045196145772934
step: 380, loss: 0.0007122300448827446
```

```

step: 390, loss: 0.00030534653342328966
step: 400, loss: 0.005096752196550369
step: 410, loss: 0.0006113793351687491
step: 420, loss: 0.00022431851539295167
step: 430, loss: 0.008875364437699318
step: 440, loss: 0.0009604044025763869
step: 450, loss: 0.0029622248839586973
step: 460, loss: 0.00020752966520376503
step: 470, loss: 0.0002838623768184334
step: 480, loss: 0.0001333390100626275
step: 490, loss: 0.006025656126439571
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 500, loss: 0.0003266028070356697
step: 510, loss: 0.016771133989095688
step: 520, loss: 0.011359707452356815
step: 530, loss: 0.036190494894981384
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9362677670793214, f1=0.9259088817303267,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015440539456903934
step: 10, loss: 0.0011044725542888045
step: 20, loss: 0.0003321695839986205
step: 30, loss: 0.005973921623080969
step: 40, loss: 0.00020897132344543934
step: 50, loss: 0.00010642661800375208
step: 60, loss: 0.009908880107104778
step: 70, loss: 0.0007165849092416465
step: 80, loss: 0.0006328907329589128
step: 90, loss: 0.0021525027696043253
step: 100, loss: 0.003646726254373789
step: 110, loss: 0.0016768831992521882
step: 120, loss: 0.001036955276504159
step: 130, loss: 0.0002127356274286285
step: 140, loss: 0.00024289725115522742
step: 150, loss: 0.0012488639913499355
step: 160, loss: 0.003928112331777811
step: 170, loss: 0.00011000486847478896
step: 180, loss: 0.0011233395198360085
step: 190, loss: 0.00029067907598800957

```

```

step: 200, loss: 0.00021353396004997194
step: 210, loss: 0.00039323794771917164
step: 220, loss: 0.00023610405332874507
step: 230, loss: 0.00017519286484457552
step: 240, loss: 0.06852412223815918
step: 250, loss: 0.0013159875525161624
step: 260, loss: 0.0018622762290760875
step: 270, loss: 0.005812837742269039
step: 280, loss: 0.008067648857831955
step: 290, loss: 0.001101729809306562
step: 300, loss: 0.0018348312005400658
step: 310, loss: 0.007047594990581274
step: 320, loss: 0.0035710069350898266
step: 330, loss: 0.0006617330946028233
step: 340, loss: 0.0009721763199195266
step: 350, loss: 0.0003900434821844101
step: 360, loss: 0.0017884951084852219
step: 370, loss: 0.0014509963802993298
step: 380, loss: 0.0007405334617942572
step: 390, loss: 0.0001289093925151974
step: 400, loss: 0.0004486065008677542
step: 410, loss: 0.00015747631550766528
step: 420, loss: 0.0002560762222856283
step: 430, loss: 6.0589380154851824e-05
step: 440, loss: 0.0001382002083119005
step: 450, loss: 0.08789735287427902
step: 460, loss: 0.0019354312680661678
step: 470, loss: 0.003078971989452839
step: 480, loss: 0.0001511320733698085
step: 490, loss: 0.0030372971668839455
step: 500, loss: 0.00038049605791457
step: 510, loss: 0.0003804028092417866
step: 520, loss: 0.00032840532367117703
step: 530, loss: 0.0006227563717402518
/usr/local/lib/python3.7/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=0.9327231121281464, f1=0.9293302540415704,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00035390406264923513
step: 10, loss: 8.379959035664797e-05
step: 20, loss: 7.30970423319377e-05
step: 30, loss: 0.0011312839342281222
step: 40, loss: 0.00013523505185730755
step: 50, loss: 0.00022601228556595743

```

step: 60, loss: 0.0005304781370796263
step: 70, loss: 0.0005904525169171393
step: 80, loss: 0.004434897564351559
step: 90, loss: 0.002092203125357628
step: 100, loss: 0.004205185454338789
step: 110, loss: 0.0003314864879939705
step: 120, loss: 0.00020576656970661134
step: 130, loss: 0.0005161768640391529
step: 140, loss: 0.00019522396905813366
step: 150, loss: 0.00037042051553726196
step: 160, loss: 0.0007233366486616433
step: 170, loss: 7.22648692317307e-05
step: 180, loss: 0.0017271980177611113
step: 190, loss: 0.00010802015458466485
step: 200, loss: 0.00013387089711613953
step: 210, loss: 0.00010832558473339304
step: 220, loss: 0.00014021170500200242
step: 230, loss: 0.003926129080355167
step: 240, loss: 0.0007356039714068174
step: 250, loss: 0.006651950068771839
step: 260, loss: 0.00017336464952677488
step: 270, loss: 0.0010453024879097939
step: 280, loss: 0.000116850329504814
step: 290, loss: 6.820773705840111e-05
step: 300, loss: 7.142317917896435e-05
step: 310, loss: 0.0033540534786880016
step: 320, loss: 0.00012023474118905142
step: 330, loss: 4.8859619710128754e-05
step: 340, loss: 0.02898534946143627
step: 350, loss: 8.595141844125465e-05
step: 360, loss: 0.00018183926295023412
step: 370, loss: 9.531990508548915e-05
step: 380, loss: 8.931040792958811e-05
step: 390, loss: 3.876482878695242e-05
step: 400, loss: 0.00022347772028297186
step: 410, loss: 0.0002144345489796251
step: 420, loss: 0.0016645106952637434
step: 430, loss: 9.826544555835426e-05
step: 440, loss: 0.00020720595784951001
step: 450, loss: 0.0001503115490777418
step: 460, loss: 0.00019554859318304807
step: 470, loss: 0.00015897002595011145
step: 480, loss: 0.0002554327074903995
step: 490, loss: 0.00021405497682280838
step: 500, loss: 0.00016864600183907896
step: 510, loss: 0.00011103812721557915
step: 520, loss: 0.0014382416848093271
step: 530, loss: 0.0004113998147659004

```

/usr/local/lib/python3.7/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=0.9285389167045972, f1=0.9258753979081401,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019888958195224404
step: 10, loss: 0.0001267176994588226
step: 20, loss: 0.0007098838104866445
step: 30, loss: 0.00025224938872270286
step: 40, loss: 0.001958019332960248
step: 50, loss: 0.00011936207010876387
step: 60, loss: 0.001714699319563806
step: 70, loss: 0.000139984447741881
step: 80, loss: 0.2414475381374359
step: 90, loss: 0.0003743777342606336
step: 100, loss: 0.03079436719417572
step: 110, loss: 0.0007942226366139948
step: 120, loss: 0.00026140132104046643
step: 130, loss: 0.0005109406192786992
step: 140, loss: 0.0065711503848433495
step: 150, loss: 0.0006559110479429364
step: 160, loss: 0.00023261918977368623
step: 170, loss: 0.0012629989068955183
step: 180, loss: 0.00018708172137849033
step: 190, loss: 0.0003163266519550234
step: 200, loss: 0.004267554730176926
step: 210, loss: 0.012384366244077682
step: 220, loss: 0.0001523498649476096
step: 230, loss: 9.437319386051968e-05
step: 240, loss: 0.0017059330129995942
step: 250, loss: 0.00011010841262759641
step: 260, loss: 0.00013970410509500653
step: 270, loss: 0.00581360561773181
step: 280, loss: 0.0005750980926677585
step: 290, loss: 6.874301470816135e-05
step: 300, loss: 0.00031703300192020833
step: 310, loss: 0.0003673978499136865
step: 320, loss: 0.002130653243511915
step: 330, loss: 0.00030832711490802467
step: 340, loss: 7.357646973105147e-05
step: 350, loss: 0.00030208821408450603
step: 360, loss: 0.02932494878768921
step: 370, loss: 0.00012550552492029965
step: 380, loss: 5.8747278671944514e-05
step: 390, loss: 0.00017904341802932322

```

```

step: 400, loss: 4.233346408000216e-05
step: 410, loss: 0.0018969988450407982
step: 420, loss: 0.00041690326179377735
step: 430, loss: 7.481456850655377e-05
step: 440, loss: 8.673613774590194e-05
step: 450, loss: 0.023318126797676086
step: 460, loss: 0.00010195787763223052
step: 470, loss: 0.018831325694918633
step: 480, loss: 0.07057234644889832
step: 490, loss: 0.0028973277658224106
step: 500, loss: 4.5241024054121226e-05
step: 510, loss: 0.0002163470780942589
step: 520, loss: 0.014123985543847084
step: 530, loss: 0.0001907268160721287
/usr/local/lib/python3.7/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=0.9315693430656934, f1=0.9337568058076225,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000496270542498678
step: 10, loss: 0.0001807117514545098
step: 20, loss: 0.00032145919976755977
step: 30, loss: 0.0001057453264365904
step: 40, loss: 0.0001950684527400881
step: 50, loss: 0.0024732989259064198
step: 60, loss: 0.00012736202916130424
step: 70, loss: 0.0007271224167197943
step: 80, loss: 0.00022832323156762868
step: 90, loss: 0.0002941954298876226
step: 100, loss: 9.697068890091032e-05
step: 110, loss: 4.002750210929662e-05
step: 120, loss: 7.999161607585847e-05
step: 130, loss: 5.9919158957200125e-05
step: 140, loss: 0.0001955080369953066
step: 150, loss: 0.0001680971181485802
step: 160, loss: 0.00022944629017729312
step: 170, loss: 0.0006074936245568097
step: 180, loss: 0.0016425808425992727
step: 190, loss: 0.00012465320469345897
step: 200, loss: 0.0025890229735523462
step: 210, loss: 0.00010616982035571709
step: 220, loss: 5.704180875909515e-05
step: 230, loss: 0.00013745957403443754
step: 240, loss: 0.0003512621915433556
step: 250, loss: 0.007265718188136816

```

```

step: 260, loss: 5.96654717810452e-05
step: 270, loss: 0.00030214007711037993
step: 280, loss: 8.039870590437204e-05
step: 290, loss: 5.795833931188099e-05
step: 300, loss: 0.0003704381815623492
step: 310, loss: 6.610876880586147e-05
step: 320, loss: 4.648297544918023e-05
step: 330, loss: 0.018533555790781975
step: 340, loss: 7.121107773855329e-05
step: 350, loss: 8.096508827293292e-05
step: 360, loss: 0.01551082544028759
step: 370, loss: 4.668417386710644e-05
step: 380, loss: 4.5398512156680226e-05
step: 390, loss: 0.00017274045967496932
step: 400, loss: 0.0001817491283873096
step: 410, loss: 0.000625659478828311
step: 420, loss: 0.0003086747310589999
step: 430, loss: 0.00011530660412972793
step: 440, loss: 4.565032941172831e-05
step: 450, loss: 9.167464304482564e-05
step: 460, loss: 0.07564923167228699
step: 470, loss: 0.001983547816053033
step: 480, loss: 2.7311085432302207e-05
step: 490, loss: 9.877424599835649e-05
step: 500, loss: 2.8871887479908764e-05
step: 510, loss: 7.411933620460331e-05
step: 520, loss: 0.00010896822641370818
step: 530, loss: 0.0002949864137917757
/usr/local/lib/python3.7/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=0.9338919925512105, f1=0.9306466729147141,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013228881871327758
step: 10, loss: 0.0002313986187800765
step: 20, loss: 5.918269380345009e-05
step: 30, loss: 0.0001139300293289125
step: 40, loss: 0.00012739682279061526
step: 50, loss: 0.1022888720035553
step: 60, loss: 8.46331677166745e-05
step: 70, loss: 8.591072401031852e-05
step: 80, loss: 0.00029852805892005563
step: 90, loss: 6.434969691326842e-05
step: 100, loss: 6.43156236037612e-05
step: 110, loss: 0.0015446225879713893

```

```
step: 120, loss: 6.0934853536309674e-05
step: 130, loss: 5.495732693816535e-05
step: 140, loss: 0.00512182991951704
step: 150, loss: 7.477123290300369e-05
step: 160, loss: 0.008753906935453415
step: 170, loss: 7.158060179790482e-05
step: 180, loss: 5.3614592616213486e-05
step: 190, loss: 7.794410339556634e-05
step: 200, loss: 0.00015942983736749738
step: 210, loss: 0.00010119451326318085
step: 220, loss: 5.7605378970038146e-05
step: 230, loss: 5.5307260481640697e-05
step: 240, loss: 0.004442162346094847
step: 250, loss: 0.0013818012084811926
step: 260, loss: 4.143272235523909e-05
step: 270, loss: 7.774063851684332e-05
step: 280, loss: 0.0001058164780260995
step: 290, loss: 3.906477286363952e-05
step: 300, loss: 0.0002236540021840483
step: 310, loss: 7.488634582841769e-05
step: 320, loss: 0.00011063790589105338
step: 330, loss: 0.0006930183153599501
step: 340, loss: 6.402521103154868e-05
step: 350, loss: 5.150521610630676e-05
step: 360, loss: 7.867306703701615e-05
step: 370, loss: 0.004141377750784159
step: 380, loss: 0.0004178885428700596
step: 390, loss: 0.00023572756617795676
step: 400, loss: 0.001089574070647359
step: 410, loss: 2.0800875063287094e-05
step: 420, loss: 4.814854764845222e-05
step: 430, loss: 0.0007572854519821703
step: 440, loss: 4.0896320570027456e-05
step: 450, loss: 0.0006041480810381472
step: 460, loss: 0.022441893815994263
step: 470, loss: 8.586591138737276e-05
step: 480, loss: 0.00014946071314625442
step: 490, loss: 0.00010925615788437426
step: 500, loss: 9.241319639841095e-05
step: 510, loss: 0.06278534978628159
step: 520, loss: 6.487542850663885e-05
step: 530, loss: 0.0007334593683481216
/usr/local/lib/python3.7/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=0.9362677670793214, f1=0.9308755760368663,
best_f1=0.9259088817303267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```


An input tensor was not cuda.

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

```
step: 0, loss: 7.266851025633514e-05
step: 10, loss: 8.272752165794373e-05
step: 20, loss: 0.0002794316387735307
step: 30, loss: 8.242473268182948e-05
step: 40, loss: 9.538480662740767e-05
step: 50, loss: 9.054369729710743e-05
step: 60, loss: 0.00019979331409558654
step: 70, loss: 0.00013891975686419755
step: 80, loss: 6.594962906092405e-05
step: 90, loss: 4.1710907680680975e-05
step: 100, loss: 3.373419167473912e-05
step: 110, loss: 0.00010918680345639586
step: 120, loss: 4.423322752700187e-05
step: 130, loss: 4.970522059011273e-05
step: 140, loss: 5.209200389799662e-05
step: 150, loss: 6.219816714292392e-05
step: 160, loss: 8.173240348696709e-05
step: 170, loss: 3.6502082366496325e-05
step: 180, loss: 0.00011171281948918477
step: 190, loss: 0.0015375040238723159
step: 200, loss: 4.287693809601478e-05
step: 210, loss: 0.00014849517901893705
step: 220, loss: 6.904893962200731e-05
step: 230, loss: 0.2222173660993576
step: 240, loss: 0.001549765351228416
step: 250, loss: 3.3192962291650474e-05
step: 260, loss: 4.245625314069912e-05
step: 270, loss: 0.00012047795462422073
step: 280, loss: 4.577423169394024e-05
step: 290, loss: 3.457821730989963e-05
step: 300, loss: 4.0851162339095026e-05
step: 310, loss: 7.121434464352205e-05
step: 320, loss: 0.0049459622241556644
step: 330, loss: 8.265711949206889e-05
step: 340, loss: 7.05920610926114e-05
step: 350, loss: 0.04684591293334961
step: 360, loss: 9.562411287333816e-05
step: 370, loss: 9.249311551684514e-05
step: 380, loss: 0.00016047077951952815
step: 390, loss: 6.239896902116016e-05
step: 400, loss: 0.00012651389988604933
step: 410, loss: 0.0008546349708922207
step: 420, loss: 4.800245733349584e-05
step: 430, loss: 2.963902079500258e-05
step: 440, loss: 0.0005451733595691621
step: 450, loss: 0.005008205305784941
```

```

step: 460, loss: 6.285251583904028e-05
step: 470, loss: 4.726892075268552e-05
step: 480, loss: 0.0010399610036984086
step: 490, loss: 6.429854693124071e-05
step: 500, loss: 0.00012040587171213701
step: 510, loss: 8.003642869880423e-05
step: 520, loss: 0.0005819337675347924
step: 530, loss: 3.7166373658692464e-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=0.9350411710887466, f1=0.9323515876668201,
best_f1=0.9259088817303267
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 [00:36, 155.54it/s]
load_f1 = 0.9365808823529411
real_f1 = 0.9353507565337
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
4398it [00:31, 138.32it/s]
```

2.1.4 Amazon-Google - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.47057685256004333
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.44022688269615173
step: 20, loss: 0.398640900850296
step: 30, loss: 0.2861962616443634
step: 40, loss: 0.3886429965496063
step: 50, loss: 0.40738990902900696
step: 60, loss: 0.40930142998695374
step: 70, loss: 0.36534908413887024
step: 80, loss: 0.4007253348827362
step: 90, loss: 0.23915813863277435
step: 100, loss: 0.24649623036384583
step: 110, loss: 0.2797473669052124
step: 120, loss: 0.3492371737957001
step: 130, loss: 0.24358147382736206
step: 140, loss: 0.47904351353645325
step: 150, loss: 0.33665376901626587
step: 160, loss: 0.4535064101219177
step: 170, loss: 0.20035699009895325
step: 180, loss: 0.2889772057533264
step: 190, loss: 0.3997660279273987
step: 200, loss: 0.2717917561531067
step: 210, loss: 0.41316720843315125
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3414634146341463, f1=0.421455938697318,
best_f1=0.421455938697318
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3301513195037842
step: 10, loss: 0.10639730095863342
step: 20, loss: 0.3723852336406708
step: 30, loss: 0.5049834251403809
step: 40, loss: 0.4261194169521332
step: 50, loss: 0.20247656106948853
step: 60, loss: 0.3137248158454895
step: 70, loss: 0.32034456729888916
step: 80, loss: 0.21514977514743805
step: 90, loss: 0.28681719303131104
step: 100, loss: 0.46909281611442566
step: 110, loss: 0.27773162722587585
step: 120, loss: 0.11756698042154312
step: 130, loss: 0.15547122061252594
step: 140, loss: 0.1047646626830101
step: 150, loss: 0.4442128539085388
step: 160, loss: 0.09338098764419556
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.38864991068840027
step: 180, loss: 0.25050365924835205
step: 190, loss: 0.3426584303379059
step: 200, loss: 0.10370062291622162
step: 210, loss: 0.2936111390590668
/usr/local/lib/python3.7/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=0.22910216718266255, f1=0.19108280254777069,
best_f1=0.421455938697318
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2673249840736389
step: 10, loss: 0.1912701278924942
step: 20, loss: 0.41207802295684814
step: 30, loss: 0.1904943585395813
step: 40, loss: 0.3420887887477875
step: 50, loss: 0.39855310320854187
step: 60, loss: 0.5102981328964233

```

```

step: 70, loss: 0.24098879098892212
step: 80, loss: 0.3690958619117737
step: 90, loss: 0.21288812160491943
step: 100, loss: 0.25694984197616577
step: 110, loss: 0.21709349751472473
step: 120, loss: 0.12235189229249954
step: 130, loss: 0.22941982746124268
step: 140, loss: 0.2192024141550064
step: 150, loss: 0.13872356712818146
step: 160, loss: 0.20644626021385193
step: 170, loss: 0.19063812494277954
step: 180, loss: 0.13821102678775787
step: 190, loss: 0.11106958985328674
step: 200, loss: 0.17342796921730042
step: 210, loss: 0.2556306719779968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4928716904276986, f1=0.496, best_f1=0.496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10709405690431595
step: 10, loss: 0.18018198013305664
step: 20, loss: 0.11947031319141388
step: 30, loss: 0.11683034151792526
step: 40, loss: 0.1918284147977829
step: 50, loss: 0.2341548651456833
step: 60, loss: 0.2620646059513092
step: 70, loss: 0.17270582914352417
step: 80, loss: 0.12131622433662415
step: 90, loss: 0.15781639516353607
step: 100, loss: 0.21975520253181458
step: 110, loss: 0.46440234780311584
step: 120, loss: 0.19759561121463776
step: 130, loss: 0.29947271943092346
step: 140, loss: 0.518985390663147
step: 150, loss: 0.2890082001686096
step: 160, loss: 0.24420596659183502
step: 170, loss: 0.2281530797481537
step: 180, loss: 0.01988767832517624
step: 190, loss: 0.25109782814979553
step: 200, loss: 0.07978051155805588
step: 210, loss: 0.3075094223022461

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5514223194748359, f1=0.501138952164009,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1652311533689499
step: 10, loss: 0.03901650011539459
step: 20, loss: 0.11579795926809311
step: 30, loss: 0.06315837800502777
step: 40, loss: 0.2048652023077011
step: 50, loss: 0.26353585720062256
step: 60, loss: 0.11511317640542984
step: 70, loss: 0.41361215710639954
step: 80, loss: 0.12631700932979584
step: 90, loss: 0.0808449238538742
step: 100, loss: 0.04091925919055939
step: 110, loss: 0.06808255612850189
step: 120, loss: 0.092070072889328
step: 130, loss: 0.12346461415290833
step: 140, loss: 0.10860294103622437
step: 150, loss: 0.18166585266590118
step: 160, loss: 0.08917096257209778
step: 170, loss: 0.33394405245780945
step: 180, loss: 0.17635808885097504
step: 190, loss: 0.1974513977766037
step: 200, loss: 0.19451329112052917
step: 210, loss: 0.1387922763824463
/usr/local/lib/python3.7/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=0.5392491467576791, f1=0.5507246376811594,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03379688411951065
step: 10, loss: 0.10711834579706192
step: 20, loss: 0.08800775557756424
step: 30, loss: 0.19060401618480682
step: 40, loss: 0.11963289976119995
step: 50, loss: 0.21711328625679016

```

```

step: 60, loss: 0.09515660256147385
step: 70, loss: 0.028228972107172012
step: 80, loss: 0.15631809830665588
step: 90, loss: 0.17689137160778046
step: 100, loss: 0.20765048265457153
step: 110, loss: 0.04456058144569397
step: 120, loss: 0.042772259563207626
step: 130, loss: 0.09509594738483429
step: 140, loss: 0.11223773658275604
step: 150, loss: 0.03343861177563667
step: 160, loss: 0.10208284854888916
step: 170, loss: 0.10856091976165771
step: 180, loss: 0.10906998068094254
step: 190, loss: 0.07345253229141235
step: 200, loss: 0.11473946273326874
step: 210, loss: 0.07323312759399414
/usr/local/lib/python3.7/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=0.5295238095238096, f1=0.5346938775510204,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12923932075500488
step: 10, loss: 0.07607994973659515
step: 20, loss: 0.038453925400972366
step: 30, loss: 0.02171974442899227
step: 40, loss: 0.006569000892341137
step: 50, loss: 0.09373444318771362
step: 60, loss: 0.1459779590368271
step: 70, loss: 0.011428261175751686
step: 80, loss: 0.1206008642911911
step: 90, loss: 0.12418156117200851
step: 100, loss: 0.12221723794937134
step: 110, loss: 0.015352763235569
step: 120, loss: 0.06289944052696228
step: 130, loss: 0.163813054561615
step: 140, loss: 0.026976367458701134
step: 150, loss: 0.05725480988621712
step: 160, loss: 0.13330133259296417
step: 170, loss: 0.07363133132457733
step: 180, loss: 0.08322969824075699
step: 190, loss: 0.016224730759859085
step: 200, loss: 0.0386897474527359
step: 210, loss: 0.14906767010688782
/usr/local/lib/python3.7/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=0.5093632958801498, f1=0.5393258426966293,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1483946442604065
step: 10, loss: 0.05223635956645012
step: 20, loss: 0.020746050402522087
step: 30, loss: 0.0355597585439682
step: 40, loss: 0.1824563890695572
step: 50, loss: 0.13441625237464905
step: 60, loss: 0.11753232032060623
step: 70, loss: 0.17182520031929016
step: 80, loss: 0.07581929117441177
step: 90, loss: 0.11523384600877762
step: 100, loss: 0.08647292852401733
step: 110, loss: 0.02665085531771183
step: 120, loss: 0.06211115047335625
step: 130, loss: 0.07920731604099274
step: 140, loss: 0.14410270750522614
step: 150, loss: 0.27377116680145264
step: 160, loss: 0.21131491661071777
step: 170, loss: 0.04429056867957115
step: 180, loss: 0.0575488805770874
step: 190, loss: 0.1420467644929886
step: 200, loss: 0.06341280788183212
step: 210, loss: 0.08957144618034363
/usr/local/lib/python3.7/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=0.5482625482625482, f1=0.5245283018867924,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05630604922771454
step: 10, loss: 0.01887933909893036
step: 20, loss: 0.02309660241007805
step: 30, loss: 0.004032495431602001
step: 40, loss: 0.002077554352581501
step: 50, loss: 0.10243944078683853
step: 60, loss: 0.018821414560079575
step: 70, loss: 0.18741054832935333
step: 80, loss: 0.0625896230340004
step: 90, loss: 0.12112221866846085
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 100, loss: 0.08218575268983841

```

```

step: 110, loss: 0.0945076197385788
step: 120, loss: 0.03928019478917122
step: 130, loss: 0.06366614997386932
step: 140, loss: 0.24015967547893524
step: 150, loss: 0.06625840812921524
step: 160, loss: 0.03350147232413292
step: 170, loss: 0.1545037180185318
step: 180, loss: 0.06227089464664459
step: 190, loss: 0.0026503868866711855
step: 200, loss: 0.007555628661066294
step: 210, loss: 0.04146617278456688
/usr/local/lib/python3.7/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=0.5419354838709678, f1=0.5296610169491526,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017531488090753555
step: 10, loss: 0.0197527464479208
step: 20, loss: 0.020056230947375298
step: 30, loss: 0.012134202755987644
step: 40, loss: 0.14515835046768188
step: 50, loss: 0.01356012187898159
step: 60, loss: 0.003949158359318972
step: 70, loss: 0.010251497849822044
step: 80, loss: 0.033448975533246994
step: 90, loss: 0.06563766300678253
step: 100, loss: 0.08967777341604233
step: 110, loss: 0.010036122053861618
step: 120, loss: 0.09615640342235565
step: 130, loss: 0.010838952846825123
step: 140, loss: 0.008103078231215477
step: 150, loss: 0.022884489968419075
step: 160, loss: 0.028789203613996506
step: 170, loss: 0.022323492914438248
step: 180, loss: 0.0240032821893692
step: 190, loss: 0.0717868059873581
step: 200, loss: 0.13190369307994843
step: 210, loss: 0.015317214652895927
/usr/local/lib/python3.7/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=0.5128205128205128, f1=0.5306122448979591,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02526705339550972
step: 10, loss: 0.023814348503947258
step: 20, loss: 0.005903994664549828
step: 30, loss: 0.017847472801804543
step: 40, loss: 0.0035406097304075956
step: 50, loss: 0.012703903019428253
step: 60, loss: 0.010120199993252754
step: 70, loss: 0.00804691482335329
step: 80, loss: 0.007026770152151585
step: 90, loss: 0.09626755118370056
step: 100, loss: 0.022044191136956215
step: 110, loss: 0.006488374900072813
step: 120, loss: 0.005074964836239815
step: 130, loss: 0.005440716166049242
step: 140, loss: 0.0029291457030922174
step: 150, loss: 0.0369836799800396
step: 160, loss: 0.000637088727671653
step: 170, loss: 0.00990633387118578
step: 180, loss: 0.013054060749709606
step: 190, loss: 0.10148080438375473
step: 200, loss: 0.003183140652254224
step: 210, loss: 0.11049973219633102
/usr/local/lib/python3.7/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=0.5135623869801085, f1=0.5369369369369369,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001324945711530745
step: 10, loss: 0.025609459728002548
step: 20, loss: 0.09105542302131653
step: 30, loss: 0.008014517836272717
step: 40, loss: 0.0003010148066096008
step: 50, loss: 0.0017354594310745597
step: 60, loss: 0.0005346747348085046
step: 70, loss: 0.0023710287641733885
step: 80, loss: 0.005288045387715101
step: 90, loss: 0.013014500960707664
step: 100, loss: 0.006034354213625193
step: 110, loss: 0.017347626388072968
step: 120, loss: 0.004246900789439678
step: 130, loss: 0.03188389539718628
step: 140, loss: 0.009081107564270496
step: 150, loss: 0.0002049483300652355
step: 160, loss: 0.022991806268692017

```

```

step: 170, loss: 0.0013255911180749536
step: 180, loss: 0.05255982652306557
step: 190, loss: 0.0010092565789818764
step: 200, loss: 0.0012571299448609352
step: 210, loss: 0.03861300274729729
/usr/local/lib/python3.7/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=0.5384615384615384, f1=0.5482041587901701,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0132059371098876
step: 10, loss: 0.022747663781046867
step: 20, loss: 0.00023941094696056098
step: 30, loss: 0.0007851270493119955
step: 40, loss: 0.008608534000813961
step: 50, loss: 0.020207200199365616
step: 60, loss: 0.08565198630094528
step: 70, loss: 0.0025555461179465055
step: 80, loss: 0.0026775833684951067
step: 90, loss: 0.005288212094455957
step: 100, loss: 0.002925093984231353
step: 110, loss: 0.04432830214500427
step: 120, loss: 0.035289835184812546
step: 130, loss: 0.005520145408809185
step: 140, loss: 0.00257607433013618
step: 150, loss: 0.0022211282048374414
step: 160, loss: 0.015271141193807125
step: 170, loss: 0.00716851931065321
step: 180, loss: 0.0018016532994806767
step: 190, loss: 0.12021900713443756
step: 200, loss: 0.011032648384571075
step: 210, loss: 0.000615430879406631
/usr/local/lib/python3.7/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=0.5107296137339056, f1=0.5305263157894737,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001067854231223464
step: 10, loss: 0.006010159850120544
step: 20, loss: 0.00631955498829484
step: 30, loss: 0.0017635003896430135
step: 40, loss: 0.0018784207059070468

```

```

step: 50, loss: 0.024431778118014336
step: 60, loss: 0.02166070230305195
step: 70, loss: 0.00031135682365857065
step: 80, loss: 0.014051804319024086
step: 90, loss: 0.0011052308836951852
step: 100, loss: 0.00021569253294728696
step: 110, loss: 0.006588434800505638
step: 120, loss: 0.0041179233230650425
step: 130, loss: 0.03028508462011814
step: 140, loss: 0.01364207360893488
step: 150, loss: 0.0039269388653337955
step: 160, loss: 0.0005800789804197848
step: 170, loss: 0.0059949494898319244
step: 180, loss: 0.001461973413825035
step: 190, loss: 0.015728704631328583
step: 200, loss: 0.02566668950021267
step: 210, loss: 0.008528560400009155
/usr/local/lib/python3.7/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=0.5235173824130879, f1=0.5237113402061856,
best_f1=0.501138952164009
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004458356183022261
step: 10, loss: 0.0460464283823967
step: 20, loss: 0.13075023889541626
step: 30, loss: 0.0013850259128957987
step: 40, loss: 0.01767604798078537
step: 50, loss: 0.0008547450415790081
step: 60, loss: 0.0013259651605039835
step: 70, loss: 0.0006349296309053898
step: 80, loss: 0.0018121832981705666
step: 90, loss: 0.001311911386437714
step: 100, loss: 0.0026618982665240765
step: 110, loss: 0.04391893371939659
step: 120, loss: 0.0018708904972299933
step: 130, loss: 0.007082193624228239
step: 140, loss: 0.002446397440508008
step: 150, loss: 0.05567275732755661
step: 160, loss: 0.021022789180278778
step: 170, loss: 0.019703274592757225
step: 180, loss: 0.0009262463427148759
step: 190, loss: 0.006678009871393442
step: 200, loss: 0.0004009292460978031
step: 210, loss: 0.00435854634270072
/usr/local/lib/python3.7/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=0.5200845665961945, f1=0.5210084033613446,
best_f1=0.501138952164009
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, 463.93it/s]
load_f1 = 0.5493562231759657
real_f1 = 0.5511111111111111
267it [00:00, 1245.56it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 203.55it/s]

```

2.1.5 Walmart-Amazon - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 619kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 35.3MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 28.4MB/s]
Downloading: 100% 501M/501M [00:10<00:00, 50.1MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type       : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4512113332748413
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.40071263909339905
step: 20, loss: 0.2611177861690521
step: 30, loss: 0.36583060026168823
step: 40, loss: 0.23873446881771088
step: 50, loss: 0.3057631850242615
step: 60, loss: 0.4374149739742279
step: 70, loss: 0.4442554712295532
step: 80, loss: 0.15269428491592407
step: 90, loss: 0.3096036911010742
step: 100, loss: 0.42504453659057617
step: 110, loss: 0.23296961188316345
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.38019582629203796
step: 130, loss: 0.3042263388633728
step: 140, loss: 0.199650838971138
step: 150, loss: 0.30961912870407104
step: 160, loss: 0.2492312788963318
step: 170, loss: 0.36791399121284485
step: 180, loss: 0.16620618104934692
step: 190, loss: 0.16056863963603973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17255252570406795, f1=0.17165847116674118,
best_f1=0.17165847116674118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41196292638778687
step: 10, loss: 0.2930580973625183
step: 20, loss: 0.6274330615997314
step: 30, loss: 0.24216675758361816
step: 40, loss: 0.563340425491333
step: 50, loss: 0.33554893732070923
step: 60, loss: 0.4800533354282379
step: 70, loss: 0.31294554471969604
step: 80, loss: 0.1561821848154068
step: 90, loss: 0.33851486444473267
step: 100, loss: 0.2613258361816406
step: 110, loss: 0.37431228160858154
step: 120, loss: 0.23577766120433807
step: 130, loss: 0.4989562928676605
step: 140, loss: 0.33381709456443787
step: 150, loss: 0.321774959564209
step: 160, loss: 0.30783531069755554
step: 170, loss: 0.23905570805072784
step: 180, loss: 0.16909779608249664
step: 190, loss: 0.23969274759292603
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17165847116674118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3942989408969879
step: 10, loss: 0.3729674220085144
step: 20, loss: 0.48591774702072144
step: 30, loss: 0.30101558566093445
step: 40, loss: 0.08963991701602936
step: 50, loss: 0.3571828305721283
step: 60, loss: 0.18615339696407318
step: 70, loss: 0.37721577286720276
step: 80, loss: 0.3133876621723175
step: 90, loss: 0.3670804798603058
step: 100, loss: 0.5308944582939148

```

```

step: 110, loss: 0.6614845991134644
step: 120, loss: 0.37457215785980225
step: 130, loss: 0.1575506031513214
step: 140, loss: 0.38978925347328186
step: 150, loss: 0.30670365691185
step: 160, loss: 0.5829534530639648
step: 170, loss: 0.4392128586769104
step: 180, loss: 0.4142517149448395
step: 190, loss: 0.16294798254966736
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17165847116674118
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2384049892425537
step: 10, loss: 0.23356537520885468
step: 20, loss: 0.29951077699661255
step: 30, loss: 0.23040369153022766
step: 40, loss: 0.5254533290863037
step: 50, loss: 0.22994045913219452
step: 60, loss: 0.3758431077003479
step: 70, loss: 0.2874874174594879
step: 80, loss: 0.215135857462883
step: 90, loss: 0.1656908243894577
step: 100, loss: 0.37237653136253357
step: 110, loss: 0.43771079182624817
step: 120, loss: 0.21496860682964325
step: 130, loss: 0.41639605164527893
step: 140, loss: 0.366245299577713
step: 150, loss: 0.22688212990760803
step: 160, loss: 0.30455127358436584
step: 170, loss: 0.38771477341651917
step: 180, loss: 0.3485363721847534
step: 190, loss: 0.14125657081604004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.1732886548587841, f1=0.17763845350052246,
best_f1=0.17763845350052246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37308764457702637
step: 10, loss: 0.37572842836380005
step: 20, loss: 0.14306342601776123
step: 30, loss: 0.12983831763267517
step: 40, loss: 0.30720996856689453
step: 50, loss: 0.5058145523071289
step: 60, loss: 0.24166560173034668
step: 70, loss: 0.37300655245780945
step: 80, loss: 0.3408999741077423
step: 90, loss: 0.2786271274089813
step: 100, loss: 0.4205271601676941
step: 110, loss: 0.3727059066295624
step: 120, loss: 0.2563682198524475
step: 130, loss: 0.5280582904815674
step: 140, loss: 0.3850794732570648
step: 150, loss: 0.3132009208202362
step: 160, loss: 0.160552516579628
step: 170, loss: 0.38917139172554016
step: 180, loss: 0.251452773809433
step: 190, loss: 0.3092745244503021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17763845350052246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30981120467185974
step: 10, loss: 0.24243669211864471
step: 20, loss: 0.3103092312812805
step: 30, loss: 0.4803442060947418
step: 40, loss: 0.2406260222196579
step: 50, loss: 0.3107726573944092
step: 60, loss: 0.4279722273349762
step: 70, loss: 0.3196614682674408
step: 80, loss: 0.31866681575775146
step: 90, loss: 0.23533140122890472
step: 100, loss: 0.4751221835613251
step: 110, loss: 0.23675936460494995
step: 120, loss: 0.44699159264564514
step: 130, loss: 0.5087694525718689
step: 140, loss: 0.194551020860672
step: 150, loss: 0.3923151195049286
step: 160, loss: 0.38509559631347656
step: 170, loss: 0.3916145861148834
step: 180, loss: 0.18637993931770325

```

```

step: 190, loss: 0.3164730668067932
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17763845350052246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.31311073899269104
step: 10, loss: 0.32056108117103577
step: 20, loss: 0.3078354001045227
step: 30, loss: 0.19377295672893524
step: 40, loss: 0.30551284551620483
step: 50, loss: 0.08096975833177567
step: 60, loss: 0.14803850650787354
step: 70, loss: 0.15400297939777374
step: 80, loss: 0.18985776603221893
step: 90, loss: 0.2403053194284439
step: 100, loss: 0.5375951528549194
step: 110, loss: 0.43999040126800537
step: 120, loss: 0.4118559658527374
step: 130, loss: 0.32220718264579773
step: 140, loss: 0.25491467118263245
step: 150, loss: 0.29289618134498596
step: 160, loss: 0.3744789958000183
step: 170, loss: 0.28721216320991516
step: 180, loss: 0.2184430956840515
step: 190, loss: 0.3049412965774536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20229633679606343, f1=0.20780648708081362,
best_f1=0.20780648708081362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21767063438892365
step: 10, loss: 0.3548811376094818
step: 20, loss: 0.20725186169147491
step: 30, loss: 0.35881686210632324
step: 40, loss: 0.152515709400177
step: 50, loss: 0.2894894778728485
step: 60, loss: 0.47860458493232727

```

```

step: 70, loss: 0.23471425473690033
step: 80, loss: 0.36408770084381104
step: 90, loss: 0.22410070896148682
step: 100, loss: 0.2864101529121399
step: 110, loss: 0.3776041567325592
step: 120, loss: 0.35757797956466675
step: 130, loss: 0.28084081411361694
step: 140, loss: 0.3624706566333771
step: 150, loss: 0.31046953797340393
step: 160, loss: 0.16010816395282745
step: 170, loss: 0.22944140434265137
step: 180, loss: 0.28984153270721436
step: 190, loss: 0.25439196825027466
/usr/local/lib/python3.7/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=0.2018448182311449, f1=0.20501635768811338,
best_f1=0.20780648708081362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22282510995864868
step: 10, loss: 0.09941113740205765
step: 20, loss: 0.3054395616054535
step: 30, loss: 0.18436011672019958
step: 40, loss: 0.30154186487197876
step: 50, loss: 0.3533356785774231
step: 60, loss: 0.37330129742622375
step: 70, loss: 0.15232525765895844
step: 80, loss: 0.3229697048664093
step: 90, loss: 0.7196939587593079
step: 100, loss: 0.36359894275665283
step: 110, loss: 0.34882014989852905
step: 120, loss: 0.5940971374511719
step: 130, loss: 0.285855233669281
step: 140, loss: 0.282419890165329
step: 150, loss: 0.23816075921058655
step: 160, loss: 0.21643340587615967
step: 170, loss: 0.5494287014007568
step: 180, loss: 0.3471270799636841
step: 190, loss: 0.2326165735721588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20579874928936895, f1=0.2165206508135169,
best_f1=0.2165206508135169
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2509251832962036
step: 10, loss: 0.1481095850467682
step: 20, loss: 0.29833483695983887
step: 30, loss: 0.3590143918991089
step: 40, loss: 0.3014131784439087
step: 50, loss: 0.10211319476366043
step: 60, loss: 0.29828599095344543
step: 70, loss: 0.3006417155265808
step: 80, loss: 0.4341222941875458
step: 90, loss: 0.1473262459039688
step: 100, loss: 0.21277201175689697
step: 110, loss: 0.21704109013080597
step: 120, loss: 0.3655395805835724
step: 130, loss: 0.4516887962818146
step: 140, loss: 0.34047019481658936
step: 150, loss: 0.2813602089881897
step: 160, loss: 0.21761293709278107
step: 170, loss: 0.3281923532485962
step: 180, loss: 0.21683788299560547
step: 190, loss: 0.14434944093227386
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2383134738771769, f1=0.24775224775224772,
best_f1=0.24775224775224772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.48076868057250977
step: 10, loss: 0.09890998154878616
step: 20, loss: 0.20699717104434967
step: 30, loss: 0.20552456378936768
step: 40, loss: 0.07632259279489517
step: 50, loss: 0.29681968688964844
step: 60, loss: 0.22563320398330688
step: 70, loss: 0.5965323448181152
step: 80, loss: 0.3749412000179291
step: 90, loss: 0.4185492694377899
step: 100, loss: 0.185311958193779

```

```

step: 110, loss: 0.2407248169183731
step: 120, loss: 0.2966213524341583
step: 130, loss: 0.6117305755615234
step: 140, loss: 0.4092196226119995
step: 150, loss: 0.2943286597728729
step: 160, loss: 0.2833141088485718
step: 170, loss: 0.4453018307685852
step: 180, loss: 0.2854449152946472
step: 190, loss: 0.12454143166542053
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.25507246376811593, f1=0.25686274509803925,
best_f1=0.25686274509803925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4243241548538208
step: 10, loss: 0.321537047624588
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.5358858108520508
step: 30, loss: 0.3522755205631256
step: 40, loss: 0.259328693151474
step: 50, loss: 0.4672923982143402
step: 60, loss: 0.3631376624107361
step: 70, loss: 0.34028881788253784
step: 80, loss: 0.3678523302078247
step: 90, loss: 0.469625324010849
step: 100, loss: 0.2759910523891449
step: 110, loss: 0.5062111020088196
step: 120, loss: 0.15774224698543549
step: 130, loss: 0.21425996720790863
step: 140, loss: 0.3288852870464325
step: 150, loss: 0.3584023714065552
step: 160, loss: 0.2124405950307846
step: 170, loss: 0.4114152491092682
step: 180, loss: 0.3020594120025635
step: 190, loss: 0.14720192551612854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2684124386252046, f1=0.3140495867768595,
best_f1=0.3140495867768595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.248131662607193
step: 10, loss: 0.26343247294425964
step: 20, loss: 0.3032388687133789
step: 30, loss: 0.3727675974369049
step: 40, loss: 0.21474666893482208
step: 50, loss: 0.21158884465694427
step: 60, loss: 0.26945558190345764
step: 70, loss: 0.1733820140361786
step: 80, loss: 0.29232776165008545
step: 90, loss: 0.28573623299598694
step: 100, loss: 0.13262434303760529
step: 110, loss: 0.595521867275238
step: 120, loss: 0.196120947599411
step: 130, loss: 0.23116834461688995
step: 140, loss: 0.29628029465675354
step: 150, loss: 0.29475143551826477
step: 160, loss: 0.42954593896865845
step: 170, loss: 0.12142058461904526
step: 180, loss: 0.186457097530365
step: 190, loss: 0.3432949185371399
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.31043956043956045, f1=0.3505747126436781,
best_f1=0.3505747126436781
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23146015405654907
step: 10, loss: 0.2083832323551178
step: 20, loss: 0.3493533730506897
step: 30, loss: 0.09459087252616882
step: 40, loss: 0.2164020985364914
step: 50, loss: 0.28364861011505127
step: 60, loss: 0.21457535028457642
step: 70, loss: 0.1976752132177353
step: 80, loss: 0.31798112392425537
step: 90, loss: 0.1726241260766983

```



```

step: 100, loss: 0.1865566074848175
step: 110, loss: 0.36334601044654846
step: 120, loss: 0.25651541352272034
step: 130, loss: 0.30272606015205383
step: 140, loss: 0.12091168761253357
step: 150, loss: 0.34863653779029846
step: 160, loss: 0.2905217409133911
step: 170, loss: 0.22475716471672058
step: 180, loss: 0.2761742174625397
step: 190, loss: 0.535416305065155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3921568627450981, f1=0.4102564102564103,
best_f1=0.4102564102564103
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.47018519043922424
step: 10, loss: 0.20225386321544647
step: 20, loss: 0.23725420236587524
step: 30, loss: 0.17884117364883423
step: 40, loss: 0.473937451839447
step: 50, loss: 0.3133050501346588
step: 60, loss: 0.14965729415416718
step: 70, loss: 0.20675021409988403
step: 80, loss: 0.2268783450126648
step: 90, loss: 0.295151948928833
step: 100, loss: 0.45167121291160583
step: 110, loss: 0.26139596104621887
step: 120, loss: 0.20947431027889252
step: 130, loss: 0.27456480264663696
step: 140, loss: 0.4655981957912445
step: 150, loss: 0.16220763325691223
step: 160, loss: 0.2859939932823181
step: 170, loss: 0.6063927412033081
step: 180, loss: 0.15568870306015015
step: 190, loss: 0.3010399043560028
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4219066937119676, f1=0.4426877470355731,
best_f1=0.4426877470355731
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:08, 235.83it/s]
load_f1 = 0.4227353463587921
real_f1 = 0.41295546558704455
733it [00:00, 3482.08it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 201.49it/s]

```

2.2 BASELINE TEXTUAL

2.2.1 Abt-Buy - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 554kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 31.0MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 26.5MB/s]
Downloading: 100% 501M/501M [00:08<00:00, 62.3MB/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: 0.4930794835090637
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5465870499610901
step: 20, loss: 0.3041968047618866
step: 30, loss: 0.4271520972251892
step: 40, loss: 0.5424747467041016
step: 50, loss: 0.31899556517601013
step: 60, loss: 0.6307167410850525
step: 70, loss: 0.352102667093277
step: 80, loss: 0.2284676432609558
step: 90, loss: 0.20688465237617493
step: 100, loss: 0.16819076240062714
step: 110, loss: 0.3658114969730377
step: 120, loss: 0.3153010308742523
step: 130, loss: 0.32658815383911133
step: 140, loss: 0.3751397728919983
step: 150, loss: 0.3064262270927429
step: 160, loss: 0.4082469344139099
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.3123278319835663
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3319411873817444
step: 10, loss: 0.47482603788375854
step: 20, loss: 0.31142446398735046
step: 30, loss: 0.3156558573246002
step: 40, loss: 0.09895572066307068
step: 50, loss: 0.5050655603408813
step: 60, loss: 0.17862574756145477
step: 70, loss: 0.5155009031295776
step: 80, loss: 0.24069133400917053
step: 90, loss: 0.25847774744033813
step: 100, loss: 0.5189167857170105
step: 110, loss: 0.2563764452934265
step: 120, loss: 0.25162872672080994
step: 130, loss: 0.5307506322860718
step: 140, loss: 0.5694282650947571
step: 150, loss: 0.44899675250053406
step: 160, loss: 0.4388158619403839
step: 170, loss: 0.39378970861434937
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.19433962264150942, f1=0.19461502125649505,
best_f1=0.19461502125649505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6319168210029602
step: 10, loss: 0.34246060252189636
step: 20, loss: 0.2429599165916443
step: 30, loss: 0.24198800325393677
step: 40, loss: 0.3670227825641632
step: 50, loss: 0.5718279480934143
step: 60, loss: 0.2973783314228058
step: 70, loss: 0.25373414158821106
step: 80, loss: 0.38310277462005615

```

```

step: 90, loss: 0.5167660713195801
step: 100, loss: 0.2784825563430786
step: 110, loss: 0.18638139963150024
step: 120, loss: 0.6058558225631714
step: 130, loss: 0.5163580775260925
step: 140, loss: 0.4552271366119385
step: 150, loss: 0.19479845464229584
step: 160, loss: 0.16712789237499237
step: 170, loss: 0.30379754304885864
/usr/local/lib/python3.7/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=0.19424799622819425, f1=0.19250829777145567,
best_f1=0.19461502125649505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37987256050109863
step: 10, loss: 0.5211026072502136
step: 20, loss: 0.2519034445285797
step: 30, loss: 0.4216921627521515
step: 40, loss: 0.24325767159461975
step: 50, loss: 0.35498490929603577
step: 60, loss: 0.700549840927124
step: 70, loss: 0.325626015663147
step: 80, loss: 0.5036882162094116
step: 90, loss: 0.2902050018310547
step: 100, loss: 0.3592243492603302
step: 110, loss: 0.42905256152153015
step: 120, loss: 0.46665289998054504
step: 130, loss: 0.3294864892959595
step: 140, loss: 0.2564072906970978
step: 150, loss: 0.7139933705329895
step: 160, loss: 0.13061437010765076
step: 170, loss: 0.2395503968000412
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24426350851221318, f1=0.23906249999999996,
best_f1=0.23906249999999996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.39351686835289

```

```

step: 10, loss: 0.2913723289966583
step: 20, loss: 0.3214244246482849
step: 30, loss: 0.33500733971595764
step: 40, loss: 0.21550512313842773
step: 50, loss: 0.24982313811779022
step: 60, loss: 0.30974894762039185
step: 70, loss: 0.3435785472393036
step: 80, loss: 0.08852926641702652
step: 90, loss: 0.5658125877380371
step: 100, loss: 0.24250949919223785
step: 110, loss: 0.25329259037971497
step: 120, loss: 0.11884021013975143
step: 130, loss: 0.2635938227176666
step: 140, loss: 0.18290597200393677
step: 150, loss: 0.30469411611557007
step: 160, loss: 0.25286728143692017
step: 170, loss: 0.35961994528770447
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2816568047337278, f1=0.23884197828709286,
best_f1=0.23884197828709286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2856124937534332
step: 10, loss: 0.12590299546718597
step: 20, loss: 0.31668737530708313
step: 30, loss: 0.22964198887348175
step: 40, loss: 0.4283159673213959
step: 50, loss: 0.23989450931549072
step: 60, loss: 0.3543631434440613
step: 70, loss: 0.3954121172428131
step: 80, loss: 0.1571207195520401
step: 90, loss: 0.32212337851524353
step: 100, loss: 0.15949013829231262
step: 110, loss: 0.4169400632381439
step: 120, loss: 0.41251179575920105
step: 130, loss: 0.5410447120666504
step: 140, loss: 0.26122069358825684
step: 150, loss: 0.18122152984142303
step: 160, loss: 0.34028202295303345
step: 170, loss: 0.2302328646183014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.308300395256917, f1=0.26631853785900783,
best_f1=0.26631853785900783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23433807492256165
step: 10, loss: 0.41663697361946106
step: 20, loss: 0.435627281665802
step: 30, loss: 0.41186732053756714
step: 40, loss: 0.10222917795181274
step: 50, loss: 0.34946802258491516
step: 60, loss: 0.4621157944202423
step: 70, loss: 0.30669522285461426
step: 80, loss: 0.12672217190265656
step: 90, loss: 0.2999856472015381
step: 100, loss: 0.17915621399879456
step: 110, loss: 0.13333512842655182
step: 120, loss: 0.21083180606365204
step: 130, loss: 0.16518352925777435
step: 140, loss: 0.1425037533044815
step: 150, loss: 0.1506483256816864
step: 160, loss: 0.047401271760463715
step: 170, loss: 0.1296653002500534
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6896551724137931, f1=0.6511627906976744,
best_f1=0.6511627906976744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2703539729118347
step: 10, loss: 0.24977515637874603
step: 20, loss: 0.13434648513793945
step: 30, loss: 0.0673731192946434
step: 40, loss: 0.10963866114616394
step: 50, loss: 0.1433248519897461
step: 60, loss: 0.1522447168827057

```



```

step: 70, loss: 0.16081586480140686
step: 80, loss: 0.23827935755252838
step: 90, loss: 0.10902538150548935
step: 100, loss: 0.052458543330430984
step: 110, loss: 0.18775109946727753
step: 120, loss: 0.37640058994293213
step: 130, loss: 0.08108887821435928
step: 140, loss: 0.3503463864326477
step: 150, loss: 0.07463674247264862
step: 160, loss: 0.08736175298690796
step: 170, loss: 0.2031695544719696
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7786666666666666, f1=0.7425474254742549,
best_f1=0.7425474254742549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15128092467784882
step: 10, loss: 0.07734446227550507
step: 20, loss: 0.045489806681871414
step: 30, loss: 0.19894611835479736
step: 40, loss: 0.22962303459644318
step: 50, loss: 0.03088223561644554
step: 60, loss: 0.12010589987039566
step: 70, loss: 0.18883028626441956
step: 80, loss: 0.13592883944511414
step: 90, loss: 0.061280373483896255
step: 100, loss: 0.11136827617883682
step: 110, loss: 0.2057483047246933
step: 120, loss: 0.0359930656850338
step: 130, loss: 0.2993497848510742
step: 140, loss: 0.043928276747465134
step: 150, loss: 0.4727402329444885
step: 160, loss: 0.06743235141038895
step: 170, loss: 0.3211849331855774
/usr/local/lib/python3.7/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=0.7727272727272727, f1=0.736842105263158,
best_f1=0.7425474254742549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07360083609819412
step: 10, loss: 0.07207586616277695
step: 20, loss: 0.1308695375919342
step: 30, loss: 0.14975605905056
step: 40, loss: 0.08203106373548508
step: 50, loss: 0.12131412327289581
step: 60, loss: 0.20789124071598053
step: 70, loss: 0.12413854897022247
step: 80, loss: 0.06138812005519867
step: 90, loss: 0.03019580990076065
step: 100, loss: 0.026760483160614967
step: 110, loss: 0.09876273572444916
step: 120, loss: 0.06803948432207108
step: 130, loss: 0.046568140387535095
step: 140, loss: 0.029278839007019997
step: 150, loss: 0.14114823937416077
step: 160, loss: 0.31442442536354065
step: 170, loss: 0.2255336344242096
/usr/local/lib/python3.7/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=0.7760416666666666, f1=0.760705289672544,
best_f1=0.7425474254742549
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3052861988544464
step: 10, loss: 0.049647919833660126
step: 20, loss: 0.02255602926015854
step: 30, loss: 0.03600144013762474
step: 40, loss: 0.013207349926233292
step: 50, loss: 0.05005945637822151
step: 60, loss: 0.09876657277345657
step: 70, loss: 0.017809752374887466
step: 80, loss: 0.012773816473782063
step: 90, loss: 0.20574939250946045
step: 100, loss: 0.2661709189414978
step: 110, loss: 0.02267805114388466
step: 120, loss: 0.031160712242126465
step: 130, loss: 0.0786866545677185
step: 140, loss: 0.346443235874176
step: 150, loss: 0.18953852355480194
step: 160, loss: 0.11803745478391647
step: 170, loss: 0.18734219670295715
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7893462469733656, f1=0.7572815533980582,
best_f1=0.7572815533980582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012392907403409481
step: 10, loss: 0.014905543997883797
step: 20, loss: 0.09237778186798096
step: 30, loss: 0.1774597316980362
step: 40, loss: 0.013741087168455124
step: 50, loss: 0.017128774896264076
step: 60, loss: 0.053332023322582245
step: 70, loss: 0.13718892633914948
step: 80, loss: 0.006649924907833338
step: 90, loss: 0.13120712339878082
step: 100, loss: 0.0150469820946455
step: 110, loss: 0.06590566784143448
step: 120, loss: 0.054296255111694336
step: 130, loss: 0.2257746458053589
step: 140, loss: 0.005527657922357321
step: 150, loss: 0.010586312972009182
step: 160, loss: 0.05860096588730812
step: 170, loss: 0.31685930490493774
/usr/local/lib/python3.7/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=0.7721179624664879, f1=0.7688311688311689,
best_f1=0.7572815533980582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009798492304980755
step: 10, loss: 0.01641228422522545
step: 20, loss: 0.00556891830638051
step: 30, loss: 0.2162536382675171
step: 40, loss: 0.11700519919395447
step: 50, loss: 0.03151160106062889
step: 60, loss: 0.009325645864009857
step: 70, loss: 0.14101450145244598
step: 80, loss: 0.011283793486654758
step: 90, loss: 0.02780883200466633
step: 100, loss: 0.07106882333755493
step: 110, loss: 0.0073404377326369286
step: 120, loss: 0.006897909566760063

```

```

step: 130, loss: 0.013493605889379978
step: 140, loss: 0.2808806002140045
step: 150, loss: 0.03608294576406479
step: 160, loss: 0.03038705699145794
step: 170, loss: 0.01168256439268589
/usr/local/lib/python3.7/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=0.7676767676767677, f1=0.768472906403941,
best_f1=0.7572815533980582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03248007595539093
step: 10, loss: 0.02727149985730648
step: 20, loss: 0.10026582330465317
step: 30, loss: 0.028816141188144684
step: 40, loss: 0.008013967424631119
step: 50, loss: 0.003460143692791462
step: 60, loss: 0.03814990818500519
step: 70, loss: 0.019264930859208107
step: 80, loss: 0.02040315978229046
step: 90, loss: 0.0059101940132677555
step: 100, loss: 0.014973842538893223
step: 110, loss: 0.0621350072324276
step: 120, loss: 0.005954614840447903
step: 130, loss: 0.12993551790714264
step: 140, loss: 0.027099302038550377
step: 150, loss: 0.01613781601190567
step: 160, loss: 0.035434942692518234
step: 170, loss: 0.08921388536691666
/usr/local/lib/python3.7/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=0.7769784172661871, f1=0.7488372093023257,
best_f1=0.7572815533980582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004207303281873465
step: 10, loss: 0.009268997237086296
step: 20, loss: 0.1984678953886032
step: 30, loss: 0.04562685266137123
step: 40, loss: 0.005938574206084013
step: 50, loss: 0.004005913157016039
step: 60, loss: 0.01380037609487772
step: 70, loss: 0.03098808228969574
step: 80, loss: 0.003453896613791585

```

```

step: 90, loss: 0.005679788067936897
step: 100, loss: 0.055714838206768036
step: 110, loss: 0.004907675553113222
step: 120, loss: 0.0062421453185379505
step: 130, loss: 0.13540972769260406
step: 140, loss: 0.005134521052241325
step: 150, loss: 0.009020286612212658
step: 160, loss: 0.18718867003917694
step: 170, loss: 0.005759597290307283
/usr/local/lib/python3.7/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=0.7745358090185676, f1=0.7751937984496124,
best_f1=0.7572815533980582
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 297.15it/s]
load_f1 = 0.7924528301886793
real_f1 = 0.7832167832167831
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:21, 203.12it/s]
```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.592879056930542
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.43627461791038513
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.46455642580986023
step: 30, loss: 0.3698261082172394
step: 40, loss: 0.3175038695335388
step: 50, loss: 0.574275016784668
step: 60, loss: 0.45768845081329346
step: 70, loss: 0.4116548001766205
step: 80, loss: 0.5340365171432495
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.43537387251853943
step: 100, loss: 0.49701282382011414
step: 110, loss: 0.4310148358345032
step: 120, loss: 0.1703328639268875
step: 130, loss: 0.2850131690502167
step: 140, loss: 0.21710236370563507
step: 150, loss: 0.1990084946155548
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 160, loss: 0.3041011691093445
step: 170, loss: 0.16995786130428314
step: 180, loss: 0.30775558948516846
step: 190, loss: 0.06788313388824463
step: 200, loss: 0.16111406683921814
step: 210, loss: 0.12177230417728424
step: 220, loss: 0.1260145753622055
step: 230, loss: 0.102922722697258
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9125799573560768, f1=0.9274725274725274,
best_f1=0.9274725274725274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1209663674235344
step: 10, loss: 0.19697800278663635
step: 20, loss: 0.09306845813989639
step: 30, loss: 0.10085217654705048
step: 40, loss: 0.06485087424516678
step: 50, loss: 0.05930695682764053
step: 60, loss: 0.056187793612480164
step: 70, loss: 0.13115271925926208
step: 80, loss: 0.02051679417490959
step: 90, loss: 0.03333275765180588
step: 100, loss: 0.03773847222328186
step: 110, loss: 0.08026367425918579
step: 120, loss: 0.10728968679904938
step: 130, loss: 0.09455772489309311
step: 140, loss: 0.018438806757330894
step: 150, loss: 0.18952874839305878
step: 160, loss: 0.042221080511808395
step: 170, loss: 0.004737692419439554
step: 180, loss: 0.054761599749326706
step: 190, loss: 0.02310527302324772
step: 200, loss: 0.035351336002349854
step: 210, loss: 0.08144903182983398
step: 220, loss: 0.14206069707870483
step: 230, loss: 0.00691111758351326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.95, f1=0.9407821229050279, best_f1=0.9407821229050279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02266438864171505
step: 10, loss: 0.07307031005620956
step: 20, loss: 0.05790058895945549
step: 30, loss: 0.03446324169635773

```



```

step: 40, loss: 0.015951650217175484
step: 50, loss: 0.05201761797070503
step: 60, loss: 0.12981246411800385
step: 70, loss: 0.028353720903396606
step: 80, loss: 0.10223980247974396
step: 90, loss: 0.06746754050254822
step: 100, loss: 0.04142234846949577
step: 110, loss: 0.02625700831413269
step: 120, loss: 0.009157312102615833
step: 130, loss: 0.033780645579099655
step: 140, loss: 0.019273225218057632
step: 150, loss: 0.19033308327198029
step: 160, loss: 0.055258966982364655
step: 170, loss: 0.008825724944472313
step: 180, loss: 0.03050648421049118
step: 190, loss: 0.05450309440493584
step: 200, loss: 0.04630909487605095
step: 210, loss: 0.003191757947206497
step: 220, loss: 0.06538532674312592
step: 230, loss: 0.11682838946580887
/usr/local/lib/python3.7/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=0.9488888888888889, f1=0.938496583143508,
best_f1=0.9407821229050279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07896661758422852
step: 10, loss: 0.05116098001599312
step: 20, loss: 0.007612364832311869
step: 30, loss: 0.0011766301468014717
step: 40, loss: 0.13804858922958374
step: 50, loss: 0.06657427549362183
step: 60, loss: 0.015766890719532967
step: 70, loss: 0.06228207051753998
step: 80, loss: 0.049534015357494354
step: 90, loss: 0.04458199813961983
step: 100, loss: 0.03924540430307388
step: 110, loss: 0.005717406515032053
step: 120, loss: 0.03487911820411682
step: 130, loss: 0.10352380573749542
step: 140, loss: 0.0264634657651186
step: 150, loss: 0.009272013790905476
step: 160, loss: 0.062294457107782364
step: 170, loss: 0.013686558231711388
step: 180, loss: 0.03356611728668213
step: 190, loss: 0.07415856420993805

```

```

step: 200, loss: 0.13344907760620117
step: 210, loss: 0.015729617327451706
step: 220, loss: 0.002083013765513897
step: 230, loss: 0.011536172591149807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9598214285714285, f1=0.9510807736063709,
best_f1=0.9510807736063709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007107822224497795
step: 10, loss: 0.024370471015572548
step: 20, loss: 0.1540871113538742
step: 30, loss: 0.03129488602280617
step: 40, loss: 0.02001381665468216
step: 50, loss: 0.06941690295934677
step: 60, loss: 0.10741329193115234
step: 70, loss: 0.008115621283650398
step: 80, loss: 0.06933657824993134
step: 90, loss: 0.10342833399772644
step: 100, loss: 0.0027654957957565784
step: 110, loss: 0.001671009580604732
step: 120, loss: 0.015259466134011745
step: 130, loss: 0.023311372846364975
step: 140, loss: 0.055546533316373825
step: 150, loss: 0.05026515573263168
step: 160, loss: 0.005679121240973473
step: 170, loss: 0.03602629527449608
step: 180, loss: 0.0061403848230838776
step: 190, loss: 0.12768006324768066
step: 200, loss: 0.1758567839860916
step: 210, loss: 0.019546281546354294
step: 220, loss: 0.009207235649228096
step: 230, loss: 0.011159537360072136
/usr/local/lib/python3.7/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=0.9576837416481068, f1=0.9525959367945823,
best_f1=0.9510807736063709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.005220579914748669
step: 10, loss: 0.008585810661315918
step: 20, loss: 0.0031307123135775328
step: 30, loss: 0.0031836931593716145
step: 40, loss: 0.0008899637614376843
step: 50, loss: 0.00478532537817955
step: 60, loss: 0.04132074490189552
step: 70, loss: 0.15395373106002808
step: 80, loss: 0.048360615968704224
step: 90, loss: 0.021066386252641678
step: 100, loss: 0.024335280060768127
step: 110, loss: 0.08961564302444458
step: 120, loss: 0.003528594272211194
step: 130, loss: 0.0035861143842339516
step: 140, loss: 0.002481528790667653
step: 150, loss: 0.0004225968150421977
step: 160, loss: 0.007493996061384678
step: 170, loss: 0.003993799909949303
step: 180, loss: 0.007574041374027729
step: 190, loss: 0.00621317932382226
step: 200, loss: 0.030809618532657623
step: 210, loss: 0.0030331166926771402
step: 220, loss: 0.007691949140280485
step: 230, loss: 0.0032664660830050707
/usr/local/lib/python3.7/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=0.9568106312292359, f1=0.9519553072625698,
best_f1=0.9510807736063709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009888955391943455
step: 10, loss: 0.0024734437465667725
step: 20, loss: 0.0010525773977860808
step: 30, loss: 0.002718496834859252
step: 40, loss: 0.14021803438663483
step: 50, loss: 0.06724955141544342
step: 60, loss: 0.0033910961356014013
step: 70, loss: 0.022597646340727806
step: 80, loss: 0.0018120544264093041
step: 90, loss: 0.00359835266135633
step: 100, loss: 0.0009404721786268055
step: 110, loss: 0.0007173437625169754
step: 120, loss: 0.006921835243701935
step: 130, loss: 0.003625878831371665
step: 140, loss: 0.0007290838402695954
step: 150, loss: 0.20092187821865082

```

```

step: 160, loss: 0.0211599450558424
step: 170, loss: 0.043090879917144775
step: 180, loss: 0.0028447939548641443
step: 190, loss: 0.13382959365844727
step: 200, loss: 0.008903201669454575
step: 210, loss: 0.023617422208189964
step: 220, loss: 0.004329224582761526
step: 230, loss: 0.002877853810787201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9621380846325166, f1=0.9430523917995445,
best_f1=0.9430523917995445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007579687051475048
step: 10, loss: 0.00327614089474082
step: 20, loss: 0.005511037539690733
step: 30, loss: 0.0018369307508692145
step: 40, loss: 0.0028001198079437017
step: 50, loss: 0.004617180209606886
step: 60, loss: 0.004439632408320904
step: 70, loss: 0.02394748106598854
step: 80, loss: 0.10449492931365967
step: 90, loss: 0.0019712685607373714
step: 100, loss: 0.0009673578315414488
step: 110, loss: 0.02023221366107464
step: 120, loss: 0.0006797181558795273
step: 130, loss: 0.00959775224328041
step: 140, loss: 0.00036682962672784925
step: 150, loss: 0.23352056741714478
step: 160, loss: 0.00053685053717345
step: 170, loss: 0.07705270498991013
step: 180, loss: 0.0027310410514473915
step: 190, loss: 0.00797877088189125
step: 200, loss: 0.019062276929616928
step: 210, loss: 0.004328195936977863
step: 220, loss: 0.0010746862972155213
step: 230, loss: 0.008528148755431175
/usr/local/lib/python3.7/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=0.96045197740113, f1=0.9437428243398392,

```

```

best_f1=0.9430523917995445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036543449386954308
step: 10, loss: 0.0007327469647862017
step: 20, loss: 0.0009375429362989962
step: 30, loss: 0.0009473390528000891
step: 40, loss: 0.01049860566854477
step: 50, loss: 0.00085030886111781
step: 60, loss: 0.000950898218434304
step: 70, loss: 0.0010256882524117827
step: 80, loss: 0.0006102052866481245
step: 90, loss: 0.05429573357105255
step: 100, loss: 0.0011027767322957516
step: 110, loss: 0.0006251267623156309
step: 120, loss: 0.18196441233158112
step: 130, loss: 0.026339493691921234
step: 140, loss: 0.07349026948213577
step: 150, loss: 0.0009740700479596853
step: 160, loss: 0.01939528062939644
step: 170, loss: 0.012964557856321335
step: 180, loss: 0.0010116015328094363
step: 190, loss: 0.000656951277051121
step: 200, loss: 0.028186185285449028
step: 210, loss: 0.016087565571069717
step: 220, loss: 0.0007268710760399699
step: 230, loss: 0.0005731579149141908
/usr/local/lib/python3.7/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=0.9610678531701891, f1=0.9457013574660633,
best_f1=0.9430523917995445
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008265148499049246
step: 10, loss: 0.0010686382884159684
step: 20, loss: 0.0011447574943304062
step: 30, loss: 0.0004689322377089411
step: 40, loss: 0.0010437901364639401
step: 50, loss: 0.0009616222814656794
step: 60, loss: 0.0007416074513457716
step: 70, loss: 0.0016604912234470248
step: 80, loss: 0.0005892579210922122
step: 90, loss: 0.0006361012347042561
step: 100, loss: 0.000915797776542604
step: 110, loss: 0.0012216351460665464

```

```

step: 120, loss: 0.0010330486111342907
step: 130, loss: 0.0032775234431028366
step: 140, loss: 0.0006048688082955778
step: 150, loss: 0.001050539780408144
step: 160, loss: 0.0004012049757875502
step: 170, loss: 0.0006933324038982391
step: 180, loss: 0.001537222764454782
step: 190, loss: 0.004176237620413303
step: 200, loss: 0.003389562014490366
step: 210, loss: 0.035964637994766235
step: 220, loss: 0.020511051639914513
step: 230, loss: 0.0006265530246309936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9655937846836848, f1=0.9492671927846673,
best_f1=0.9492671927846673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026884363032877445
step: 10, loss: 0.0017358913319185376
step: 20, loss: 0.001445063971914351
step: 30, loss: 0.008525797165930271
step: 40, loss: 0.0004979819641448557
step: 50, loss: 0.0013993083266541362
step: 60, loss: 0.02220565639436245
step: 70, loss: 0.0006760578835383058
step: 80, loss: 0.0006706491112709045
step: 90, loss: 0.19165180623531342
step: 100, loss: 0.0008439146913588047
step: 110, loss: 0.0006271966267377138
step: 120, loss: 0.0006771594053134322
step: 130, loss: 0.0006569123943336308
step: 140, loss: 0.003684178227558732
step: 150, loss: 0.000960711797233671
step: 160, loss: 0.015299015678465366
step: 170, loss: 0.0018104196060448885
step: 180, loss: 0.001594672561623156
step: 190, loss: 0.0006452499073930085
step: 200, loss: 0.02059878595173359
step: 210, loss: 0.0009735642815940082
step: 220, loss: 0.0009002229198813438
step: 230, loss: 0.0007258497644215822

```

```

/usr/local/lib/python3.7/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=0.9627959413754228, f1=0.9507445589919816,
best_f1=0.9492671927846673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015131074003875256
step: 10, loss: 0.0006001396104693413
step: 20, loss: 0.001058118650689721
step: 30, loss: 0.2841537892818451
step: 40, loss: 0.0008489011088386178
step: 50, loss: 0.0008632770041003823
step: 60, loss: 0.03846285864710808
step: 70, loss: 0.0008229077211581171
step: 80, loss: 0.00039693573489785194
step: 90, loss: 0.0007572612375952303
step: 100, loss: 0.0004951974260620773
step: 110, loss: 0.00041254505049437284
step: 120, loss: 0.009708481840789318
step: 130, loss: 0.0006487150676548481
step: 140, loss: 0.0007136339554563165
step: 150, loss: 0.0007054531597532332
step: 160, loss: 0.0005539966514334083
step: 170, loss: 0.002191950101405382
step: 180, loss: 0.00041237223194912076
step: 190, loss: 0.0007265431340783834
step: 200, loss: 0.00046207275590859354
step: 210, loss: 0.13261090219020844
step: 220, loss: 0.003418172476813197
step: 230, loss: 0.000801107962615788
/usr/local/lib/python3.7/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=0.9640449438202248, f1=0.9532497149372862,
best_f1=0.9492671927846673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007656426634639502
step: 10, loss: 0.009713065810501575
step: 20, loss: 0.0011504535796120763
step: 30, loss: 0.013928421773016453
step: 40, loss: 0.0013388519873842597
step: 50, loss: 0.007528219372034073
step: 60, loss: 0.0006769471219740808
step: 70, loss: 0.0004942219820804894

```

```

step: 80, loss: 0.002066671149805188
step: 90, loss: 0.0005243937484920025
step: 100, loss: 0.0009398983092978597
step: 110, loss: 0.0009396012756042182
step: 120, loss: 0.0007625535945408046
step: 130, loss: 0.000908120593521744
step: 140, loss: 0.0006220921059139073
step: 150, loss: 0.000362495455192402
step: 160, loss: 0.003280120901763439
step: 170, loss: 0.0009286498534493148
step: 180, loss: 0.0698290541768074
step: 190, loss: 0.0033097248524427414
step: 200, loss: 0.00034365494502708316
step: 210, loss: 0.0006833134102635086
step: 220, loss: 0.0007661546696908772
step: 230, loss: 0.04749785736203194
/usr/local/lib/python3.7/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=0.961625282167043, f1=0.944700460829493,
best_f1=0.9492671927846673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005059733521193266
step: 10, loss: 0.0004890129785053432
step: 20, loss: 0.0035524037666618824
step: 30, loss: 0.0009642249206081033
step: 40, loss: 0.00509807700291276
step: 50, loss: 0.00047045049723237753
step: 60, loss: 0.0005756297032348812
step: 70, loss: 0.00120927847456187
step: 80, loss: 0.000669394270516932
step: 90, loss: 0.0010335644474253058
step: 100, loss: 0.0008057214436121285
step: 110, loss: 0.001235164818353951
step: 120, loss: 0.0003220764338038862
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 130, loss: 0.030344510450959206
step: 140, loss: 0.0017856867052614689
step: 150, loss: 0.00038350350223481655
step: 160, loss: 0.0008228135411627591
step: 170, loss: 0.0007637323578819633
step: 180, loss: 0.0009400478447787464
step: 190, loss: 0.000540971290320158
step: 200, loss: 0.000732544285710901
step: 210, loss: 0.001866013859398663
step: 220, loss: 0.0006592720164917409

```



```

step: 230, loss: 0.0003657522320281714
/usr/local/lib/python3.7/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=0.9596412556053813, f1=0.9500000000000001,
best_f1=0.9492671927846673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03902708739042282
step: 10, loss: 0.0006829540943726897
step: 20, loss: 0.000647092645522207
step: 30, loss: 0.0007759995060041547
step: 40, loss: 0.0005279112956486642
step: 50, loss: 0.0006300462991930544
step: 60, loss: 0.00042169223888777196
step: 70, loss: 0.002220324007794261
step: 80, loss: 0.0006508767837658525
step: 90, loss: 0.0005003067199140787
step: 100, loss: 0.00043166460818611085
step: 110, loss: 0.0007152455509640276
step: 120, loss: 0.007300327531993389
step: 130, loss: 0.0004408300737850368
step: 140, loss: 0.0007146683637984097
step: 150, loss: 0.0005443163099698722
step: 160, loss: 0.0016515852184966207
step: 170, loss: 0.0003940412134397775
step: 180, loss: 0.0006647692061960697
step: 190, loss: 0.0003189428534824401
step: 200, loss: 0.0005429005832411349
step: 210, loss: 0.0015174847794696689
step: 220, loss: 0.000619755417574197
step: 230, loss: 0.0005951581988483667
/usr/local/lib/python3.7/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=0.9596412556053813, f1=0.9501133786848073,
best_f1=0.9492671927846673
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

```

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

```

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 204.87it/s]
load_f1 = 0.9675977653631285
real_f1 = 0.9634551495016611
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 204.84it/s]

```

2.3.2 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6273968815803528
```

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

```
step: 10, loss: 0.4277028441429138
```

```
step: 20, loss: 0.30690792202949524
```

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

```
step: 30, loss: 0.3865823745727539
```

```
step: 40, loss: 0.38493210077285767
```

```
step: 50, loss: 0.6546292901039124
```

```
step: 60, loss: 0.3613128662109375
```

step: 70, loss: 0.4594770669937134
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 80, loss: 0.5001224279403687
step: 90, loss: 0.42506495118141174
step: 100, loss: 0.6016262769699097
step: 110, loss: 0.38430672883987427
step: 120, loss: 0.5205445289611816
step: 130, loss: 0.5045889019966125
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 140, loss: 0.5132622718811035
step: 150, loss: 0.25808602571487427
step: 160, loss: 0.38461485505104065
step: 170, loss: 0.30832457542419434
step: 180, loss: 0.3854077458381653
step: 190, loss: 0.5565117001533508
step: 200, loss: 0.21801310777664185
step: 210, loss: 0.12024745345115662
step: 220, loss: 0.07813156396150589
step: 230, loss: 0.33988863229751587
step: 240, loss: 0.08414514362812042
step: 250, loss: 0.12494271248579025
step: 260, loss: 0.8877171874046326
step: 270, loss: 0.297230988740921
step: 280, loss: 0.150853231549263
step: 290, loss: 0.2601287066936493
step: 300, loss: 0.17560748755931854
step: 310, loss: 0.4456561207771301
step: 320, loss: 0.17006517946720123
step: 330, loss: 0.25612014532089233
step: 340, loss: 0.3742695152759552
step: 350, loss: 0.27153530716896057
step: 360, loss: 0.03691401332616806
step: 370, loss: 0.06872732937335968
step: 380, loss: 0.17096033692359924
step: 390, loss: 0.09864038974046707
step: 400, loss: 0.0847342386841774
step: 410, loss: 0.2514558434486389
step: 420, loss: 0.05040498450398445
step: 430, loss: 0.05722483992576599
step: 440, loss: 0.17048269510269165
step: 450, loss: 0.23815299570560455
step: 460, loss: 0.14863677322864532
step: 470, loss: 0.06937512010335922
step: 480, loss: 0.2537558078765869
step: 490, loss: 0.07214799523353577
step: 500, loss: 0.1015365868806839
step: 510, loss: 0.274165540933609
step: 520, loss: 0.2625181972980499

```

step: 530, loss: 0.0879729688167572
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8794664125774178, f1=0.8690307328605201,
best_f1=0.8690307328605201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05890211835503578
step: 10, loss: 0.16783717274665833
step: 20, loss: 0.17242546379566193
step: 30, loss: 0.11628584563732147
step: 40, loss: 0.22164392471313477
step: 50, loss: 0.15056288242340088
step: 60, loss: 0.1482192724943161
step: 70, loss: 0.05494583398103714
step: 80, loss: 0.0599275566637516
step: 90, loss: 0.022138679400086403
step: 100, loss: 0.07204565405845642
step: 110, loss: 0.14438611268997192
step: 120, loss: 0.18032871186733246
step: 130, loss: 0.022524511441588402
step: 140, loss: 0.12828654050827026
step: 150, loss: 0.10393574088811874
step: 160, loss: 0.08751347661018372
step: 170, loss: 0.11085531860589981
step: 180, loss: 0.1551351696252823
step: 190, loss: 0.07869629561901093
step: 200, loss: 0.27983155846595764
step: 210, loss: 0.06612598150968552
step: 220, loss: 0.01234554685652256
step: 230, loss: 0.10941710323095322
step: 240, loss: 0.07040606439113617
step: 250, loss: 0.05283575505018234
step: 260, loss: 0.053673114627599716
step: 270, loss: 0.046507351100444794
step: 280, loss: 0.061810072511434555
step: 290, loss: 0.07985055446624756
step: 300, loss: 0.14194288849830627
step: 310, loss: 0.07039877027273178
step: 320, loss: 0.11663883179426193
step: 330, loss: 0.06628741323947906
step: 340, loss: 0.22425949573516846

```

```

step: 350, loss: 0.006868560798466206
step: 360, loss: 0.1184544786810875
step: 370, loss: 0.1488162875175476
step: 380, loss: 0.19384439289569855
step: 390, loss: 0.024978918954730034
step: 400, loss: 0.0971783772110939
step: 410, loss: 0.09426165372133255
step: 420, loss: 0.1604440063238144
step: 430, loss: 0.045853909105062485
step: 440, loss: 0.006516596302390099
step: 450, loss: 0.05986351892352104
step: 460, loss: 0.17512844502925873
step: 470, loss: 0.1399141401052475
step: 480, loss: 0.04063142463564873
step: 490, loss: 0.2339000552892685
step: 500, loss: 0.012390629388391972
step: 510, loss: 0.0845717117190361
step: 520, loss: 0.30892908573150635
step: 530, loss: 0.13258977234363556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9124886052871468, f1=0.8990406578346277,
best_f1=0.8990406578346277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1417836993932724
step: 10, loss: 0.19677746295928955
step: 20, loss: 0.09766430407762527
step: 30, loss: 0.14741921424865723
step: 40, loss: 0.059805724769830704
step: 50, loss: 0.10147634148597717
step: 60, loss: 0.04385441914200783
step: 70, loss: 0.05103740096092224
step: 80, loss: 0.10563507676124573
step: 90, loss: 0.05697248503565788
step: 100, loss: 0.2005024403333664
step: 110, loss: 0.07652401179075241
step: 120, loss: 0.15590450167655945
step: 130, loss: 0.06385517865419388
step: 140, loss: 0.04057922214269638
step: 150, loss: 0.09728733450174332
step: 160, loss: 0.10428363084793091

```

```

step: 170, loss: 0.014743112027645111
step: 180, loss: 0.038581281900405884
step: 190, loss: 0.06715650111436844
step: 200, loss: 0.07715366035699844
step: 210, loss: 0.13392551243305206
step: 220, loss: 0.06868532299995422
step: 230, loss: 0.05265286937355995
step: 240, loss: 0.071482352912426
step: 250, loss: 0.3547905683517456
step: 260, loss: 0.0803355947136879
step: 270, loss: 0.04282333701848984
step: 280, loss: 0.11539648473262787
step: 290, loss: 0.025491071864962578
step: 300, loss: 0.1563318818807602
step: 310, loss: 0.2751661241054535
step: 320, loss: 0.030590562149882317
step: 330, loss: 0.02115909568965435
step: 340, loss: 0.029869215562939644
step: 350, loss: 0.23751112818717957
step: 360, loss: 0.12036118656396866
step: 370, loss: 0.07746583968400955
step: 380, loss: 0.021137477830052376
step: 390, loss: 0.09883896261453629
step: 400, loss: 0.13438774645328522
step: 410, loss: 0.05071466416120529
step: 420, loss: 0.18589460849761963
step: 430, loss: 0.090882308781147
step: 440, loss: 0.21688221395015717
step: 450, loss: 0.18276487290859222
step: 460, loss: 0.1966874748468399
step: 470, loss: 0.0738791972398758
step: 480, loss: 0.10392347723245621
step: 490, loss: 0.0433230847120285
step: 500, loss: 0.021681295707821846
step: 510, loss: 0.1143827885389328
step: 520, loss: 0.030864842236042023
step: 530, loss: 0.18638849258422852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9210770659238626, f1=0.9041608228143994,
best_f1=0.9041608228143994
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0343252494931221
step: 10, loss: 0.025840122252702713
step: 20, loss: 0.0926009938120842
step: 30, loss: 0.12920445203781128
step: 40, loss: 0.01645660772919655
step: 50, loss: 0.03377489373087883
step: 60, loss: 0.03887239471077919
step: 70, loss: 0.08616064488887787
step: 80, loss: 0.24517561495304108
step: 90, loss: 0.027489297091960907
step: 100, loss: 0.004995354451239109
step: 110, loss: 0.33801698684692383
step: 120, loss: 0.0711163729429245
step: 130, loss: 0.14571383595466614
step: 140, loss: 0.09497997164726257
step: 150, loss: 0.015026104636490345
step: 160, loss: 0.08121898025274277
step: 170, loss: 0.027238773182034492
step: 180, loss: 0.07027540355920792
step: 190, loss: 0.059289876371622086
step: 200, loss: 0.04671529680490494
step: 210, loss: 0.00764441629871726
step: 220, loss: 0.009744751267135143
step: 230, loss: 0.05662692338228226
step: 240, loss: 0.011057784780859947
step: 250, loss: 0.12087734788656235
step: 260, loss: 0.05101603642106056
step: 270, loss: 0.21617665886878967
step: 280, loss: 0.1129024475812912
step: 290, loss: 0.0679287537932396
step: 300, loss: 0.009224290028214455
step: 310, loss: 0.01254588458687067
step: 320, loss: 0.1596175730228424
step: 330, loss: 0.012909499928355217
step: 340, loss: 0.032188352197408676
step: 350, loss: 0.2124362736940384
step: 360, loss: 0.04170601814985275
step: 370, loss: 0.0077374167740345
step: 380, loss: 0.17935003340244293
step: 390, loss: 0.003825494786724448
step: 400, loss: 0.015005689114332199
step: 410, loss: 0.007756764534860849
step: 420, loss: 0.01787135750055313
step: 430, loss: 0.0072056301869452
step: 440, loss: 0.012457245029509068
step: 450, loss: 0.09588806331157684
step: 460, loss: 0.04726782813668251
```



```

step: 470, loss: 0.006242056377232075
step: 480, loss: 0.0818096324801445
step: 490, loss: 0.03624546155333519
step: 500, loss: 0.18084852397441864
step: 510, loss: 0.1311032772064209
step: 520, loss: 0.057056132704019547
step: 530, loss: 0.08544886857271194
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9298085688240657, f1=0.9153952843273233,
best_f1=0.9153952843273233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006512920837849379
step: 10, loss: 0.07485006749629974
step: 20, loss: 0.012619705870747566
step: 30, loss: 0.11247479915618896
step: 40, loss: 0.003483563894405961
step: 50, loss: 0.07042382657527924
step: 60, loss: 0.08021953701972961
step: 70, loss: 0.01907760091125965
step: 80, loss: 0.055560141801834106
step: 90, loss: 0.06403230875730515
step: 100, loss: 0.09718898683786392
step: 110, loss: 0.055009566247463226
step: 120, loss: 0.13827846944332123
step: 130, loss: 0.06085829436779022
step: 140, loss: 0.03460550680756569
step: 150, loss: 0.03582683950662613
step: 160, loss: 0.021976064890623093
step: 170, loss: 0.11760387569665909
step: 180, loss: 0.005064898636192083
step: 190, loss: 0.015185498632490635
step: 200, loss: 0.017122406512498856
step: 210, loss: 0.009869802743196487
step: 220, loss: 0.06885411590337753
step: 230, loss: 0.019375089555978775
step: 240, loss: 0.07061906158924103
step: 250, loss: 0.07806318998336792
step: 260, loss: 0.001682321191765368
step: 270, loss: 0.0026501694228500128
step: 280, loss: 0.009077226743102074

```

```

step: 290, loss: 0.008082564920186996
step: 300, loss: 0.4514341950416565
step: 310, loss: 0.06596847623586655
step: 320, loss: 0.11772923171520233
step: 330, loss: 0.015730755403637886
step: 340, loss: 0.017604323104023933
step: 350, loss: 0.003410415956750512
step: 360, loss: 0.0018047827761620283
step: 370, loss: 0.008966137655079365
step: 380, loss: 0.0014107389142736793
step: 390, loss: 0.21927565336227417
step: 400, loss: 0.005263683386147022
step: 410, loss: 0.03113688714802265
step: 420, loss: 0.21255135536193848
step: 430, loss: 0.16176865994930267
step: 440, loss: 0.06806915253400803
step: 450, loss: 0.10212826728820801
step: 460, loss: 0.04823039099574089
step: 470, loss: 0.015246845781803131
step: 480, loss: 0.027549970895051956
step: 490, loss: 0.05229967460036278
step: 500, loss: 0.07505322247743607
step: 510, loss: 0.029281537979841232
step: 520, loss: 0.0400499664247036
step: 530, loss: 0.011475740000605583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9305108145421077, f1=0.9133271202236719,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04072171449661255
step: 10, loss: 0.002613755175843835
step: 20, loss: 0.003651437582448125
step: 30, loss: 0.031300704926252365
step: 40, loss: 0.03108053095638752
step: 50, loss: 0.00285562826320529
step: 60, loss: 0.0662572905421257
step: 70, loss: 0.005745571572333574
step: 80, loss: 0.00980030931532383
step: 90, loss: 0.16643521189689636
step: 100, loss: 0.006790728308260441

```

```
step: 110, loss: 0.027407104149460793
step: 120, loss: 0.03628314658999443
step: 130, loss: 0.012561457231640816
step: 140, loss: 0.07167254388332367
step: 150, loss: 0.006239232141524553
step: 160, loss: 0.09554939717054367
step: 170, loss: 0.01945933699607849
step: 180, loss: 0.0032601128332316875
step: 190, loss: 0.10369851440191269
step: 200, loss: 0.014891011640429497
step: 210, loss: 0.14321835339069366
step: 220, loss: 0.02253834903240204
step: 230, loss: 0.09591241925954819
step: 240, loss: 0.11818729341030121
step: 250, loss: 0.011669296771287918
step: 260, loss: 0.03267413750290871
step: 270, loss: 0.012273915112018585
step: 280, loss: 0.05793905630707741
step: 290, loss: 0.029719578102231026
step: 300, loss: 0.03405291214585304
step: 310, loss: 0.17845018208026886
step: 320, loss: 0.03360734134912491
step: 330, loss: 0.05453015863895416
step: 340, loss: 0.002680526813492179
step: 350, loss: 0.0036713522858917713
step: 360, loss: 0.05757157504558563
step: 370, loss: 0.004087959881871939
step: 380, loss: 0.0157973300665617
step: 390, loss: 0.025217000395059586
step: 400, loss: 0.04335714504122734
step: 410, loss: 0.0020084979478269815
step: 420, loss: 0.008605523966252804
step: 430, loss: 0.0017680602613836527
step: 440, loss: 0.014025435782968998
step: 450, loss: 0.36441943049430847
step: 460, loss: 0.013627860695123672
step: 470, loss: 0.0038959323428571224
step: 480, loss: 0.035882893949747086
step: 490, loss: 0.192092165350914
step: 500, loss: 0.04393267631530762
step: 510, loss: 0.03881125897169113
step: 520, loss: 0.005637420807033777
step: 530, loss: 0.008042127825319767
/usr/local/lib/python3.7/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=0.922863741339492, f1=0.9086672879776329,
best_f1=0.9133271202236719
```

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

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

```
step: 0, loss: 0.007150425110012293  
step: 10, loss: 0.002258060732856393  
step: 20, loss: 0.008458222262561321  
step: 30, loss: 0.029740657657384872  
step: 40, loss: 0.013197189196944237  
step: 50, loss: 0.07017440348863602  
step: 60, loss: 0.007965533062815666  
step: 70, loss: 0.004920929204672575  
step: 80, loss: 0.004362739156931639  
step: 90, loss: 0.0022340998984873295  
step: 100, loss: 0.016634397208690643  
step: 110, loss: 0.0073596579022705555  
step: 120, loss: 0.019904036074876785  
step: 130, loss: 0.0007860968471504748  
step: 140, loss: 0.002009795745834708  
step: 150, loss: 0.0018645233940333128  
step: 160, loss: 0.02694987691938877  
step: 170, loss: 0.006688939873129129  
step: 180, loss: 0.06260187923908234  
step: 190, loss: 0.009595640003681183  
step: 200, loss: 0.017119908705353737  
step: 210, loss: 0.006892874836921692  
step: 220, loss: 0.00332638225518167  
step: 230, loss: 0.016926521435379982  
step: 240, loss: 0.10549873858690262  
step: 250, loss: 0.0630280002951622  
step: 260, loss: 0.01793079823255539  
step: 270, loss: 0.010497910901904106  
step: 280, loss: 0.014983047731220722  
step: 290, loss: 0.007232113741338253  
step: 300, loss: 0.00431145029142499  
step: 310, loss: 0.014432324096560478  
step: 320, loss: 0.0067541575990617275  
step: 330, loss: 0.031591352075338364  
step: 340, loss: 0.012230297550559044  
step: 350, loss: 0.004336215555667877  
step: 360, loss: 0.002805825090035796  
step: 370, loss: 0.010532137006521225  
step: 380, loss: 0.00837814249098301  
step: 390, loss: 0.027849435806274414  
step: 400, loss: 0.03743501007556915  
step: 410, loss: 0.009145201183855534  
step: 420, loss: 0.08199642598628998  
step: 430, loss: 0.0007501025684177876  
step: 440, loss: 0.002103498438373208
```

```
step: 450, loss: 0.0907483771443367
step: 460, loss: 0.005270335357636213
step: 470, loss: 0.29900893568992615
step: 480, loss: 0.004628514871001244
step: 490, loss: 0.13538673520088196
step: 500, loss: 0.0035297716967761517
step: 510, loss: 0.004224393982440233
step: 520, loss: 0.007700163405388594
step: 530, loss: 0.0017062602564692497
/usr/local/lib/python3.7/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=0.9278350515463917, f1=0.9069437883797827,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018400212284177542
step: 10, loss: 0.006032527890056372
step: 20, loss: 0.001793667208403349
step: 30, loss: 0.038626980036497116
step: 40, loss: 0.0009274072363041341
step: 50, loss: 0.001059786882251501
step: 60, loss: 0.0013440982438623905
step: 70, loss: 0.038506947457790375
step: 80, loss: 0.000578170467633754
step: 90, loss: 0.03905090317130089
step: 100, loss: 0.008369258604943752
step: 110, loss: 0.0043771169148385525
step: 120, loss: 0.007164048030972481
step: 130, loss: 0.053053054958581924
step: 140, loss: 0.0004902890650555491
step: 150, loss: 0.0016010603867471218
step: 160, loss: 0.04135638102889061
step: 170, loss: 0.13189579546451569
step: 180, loss: 0.004271500278264284
step: 190, loss: 0.012428775429725647
step: 200, loss: 0.019006572663784027
step: 210, loss: 0.22801309823989868
step: 220, loss: 0.0034396613482385874
step: 230, loss: 0.22438064217567444
step: 240, loss: 0.007971297018229961
step: 250, loss: 0.02428949996829033
step: 260, loss: 0.0021077285055071115
step: 270, loss: 0.04805154353380203
step: 280, loss: 0.008257873356342316
step: 290, loss: 0.024312948808073997
step: 300, loss: 0.001733680721372366
```

```

step: 310, loss: 0.0020563025027513504
step: 320, loss: 0.009106753394007683
step: 330, loss: 0.0008633667603135109
step: 340, loss: 0.013598502613604069
step: 350, loss: 0.0004357954312581569
step: 360, loss: 0.03348921611905098
step: 370, loss: 0.061668865382671356
step: 380, loss: 0.0010789852822199464
step: 390, loss: 0.03457681089639664
step: 400, loss: 0.00570266367867589
step: 410, loss: 0.006620282772928476
step: 420, loss: 0.002685744082555175
step: 430, loss: 0.025347448885440826
step: 440, loss: 0.007383698131889105
step: 450, loss: 0.05626385658979416
step: 460, loss: 0.014503154903650284
step: 470, loss: 0.08740068972110748
step: 480, loss: 0.015958277508616447
step: 490, loss: 0.023322992026805878
step: 500, loss: 0.13605313003063202
step: 510, loss: 0.1462005227804184
step: 520, loss: 0.0029063241090625525
step: 530, loss: 0.022007741034030914
/usr/local/lib/python3.7/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=0.9297501178689297, f1=0.9100478468899522,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007910503190942109
step: 10, loss: 0.008323208428919315
step: 20, loss: 0.0024536524433642626
step: 30, loss: 0.06249842420220375
step: 40, loss: 0.00389784830622375
step: 50, loss: 0.0035672097001224756
step: 60, loss: 0.0028799984138458967
step: 70, loss: 0.0723731517791748
step: 80, loss: 0.014708256348967552
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.08189558982849121
step: 100, loss: 0.050670161843299866
step: 110, loss: 0.06944870203733444
step: 120, loss: 0.008479436859488487
step: 130, loss: 0.0026563939172774553
step: 140, loss: 0.0363367535173893
step: 150, loss: 0.010161385871469975

```

```

step: 160, loss: 0.01591724343597889
step: 170, loss: 0.02555600181221962
step: 180, loss: 0.014264506287872791
step: 190, loss: 0.010196978226304054
step: 200, loss: 0.0006760972901247442
step: 210, loss: 0.03216733783483505
step: 220, loss: 0.0009619120974093676
step: 230, loss: 0.013963832519948483
step: 240, loss: 0.0030114201363176107
step: 250, loss: 0.004639694932848215
step: 260, loss: 0.021875828504562378
step: 270, loss: 0.006717166863381863
step: 280, loss: 0.04764091968536377
step: 290, loss: 0.002427492057904601
step: 300, loss: 0.0038963346742093563
step: 310, loss: 0.04926535114645958
step: 320, loss: 0.0005975901731289923
step: 330, loss: 0.0022135069593787193
step: 340, loss: 0.001477596233598888
step: 350, loss: 0.036871157586574554
step: 360, loss: 0.0008345263777300715
step: 370, loss: 0.011165001429617405
step: 380, loss: 0.0013423284981399775
step: 390, loss: 0.00042644268251024187
step: 400, loss: 0.014098198153078556
step: 410, loss: 0.003836600808426738
step: 420, loss: 0.0006516557186841965
step: 430, loss: 0.0020847534760832787
step: 440, loss: 0.0050431108102202415
step: 450, loss: 0.20622530579566956
step: 460, loss: 0.10929320007562637
step: 470, loss: 0.0011727282544597983
step: 480, loss: 0.0004913975717499852
step: 490, loss: 0.0015033322852104902
step: 500, loss: 0.003516431199386716
step: 510, loss: 0.007507513742893934
step: 520, loss: 0.07423471659421921
step: 530, loss: 0.034351006150245667
/usr/local/lib/python3.7/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=0.9253456221198156, f1=0.9153810191678354,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08205285668373108
step: 10, loss: 0.0007496149046346545

```

step: 20, loss: 0.0018846483435481787
step: 30, loss: 0.005359150469303131
step: 40, loss: 0.0051935045048594475
step: 50, loss: 0.007860398851335049
step: 60, loss: 0.0020899176597595215
step: 70, loss: 0.001456160913221538
step: 80, loss: 0.00257049105130136
step: 90, loss: 0.0009562387713231146
step: 100, loss: 0.0012211394496262074
step: 110, loss: 0.0060219489969313145
step: 120, loss: 0.0005772248841822147
step: 130, loss: 0.000851387856528163
step: 140, loss: 0.0014366492396220565
step: 150, loss: 0.008472184650599957
step: 160, loss: 0.07897371053695679
step: 170, loss: 0.0012771107722073793
step: 180, loss: 0.006176554132252932
step: 190, loss: 0.0022264837753027678
step: 200, loss: 0.01999930664896965
step: 210, loss: 0.01807377301156521
step: 220, loss: 0.0013547297567129135
step: 230, loss: 0.0009112640982493758
step: 240, loss: 0.02804764173924923
step: 250, loss: 0.0023353183642029762
step: 260, loss: 0.1981593370437622
step: 270, loss: 0.006700091063976288
step: 280, loss: 0.005551077891141176
step: 290, loss: 0.0008607865311205387
step: 300, loss: 0.0012826549354940653
step: 310, loss: 0.007766381837427616
step: 320, loss: 0.005559214390814304
step: 330, loss: 0.0023080697283148766
step: 340, loss: 0.0008909482276067138
step: 350, loss: 0.0651603639125824
step: 360, loss: 0.0009260232327505946
step: 370, loss: 0.003235175274312496
step: 380, loss: 0.004473415203392506
step: 390, loss: 0.000538280641194433
step: 400, loss: 0.05616549029946327
step: 410, loss: 0.002129905391484499
step: 420, loss: 0.000691277498845011
step: 430, loss: 0.0022835314739495516
step: 440, loss: 0.0024196822196245193
step: 450, loss: 0.000713749963324517
step: 460, loss: 0.0009362497949041426
step: 470, loss: 0.08093162626028061
step: 480, loss: 0.005087852478027344
step: 490, loss: 0.0008162165759131312


```

step: 500, loss: 0.006211560219526291
step: 510, loss: 0.0024712313897907734
step: 520, loss: 0.0006742349360138178
step: 530, loss: 0.002423943020403385
/usr/local/lib/python3.7/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=0.9274004683840749, f1=0.9166274140367404,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001366847543977201
step: 10, loss: 0.006323530804365873
step: 20, loss: 0.010915164835751057
step: 30, loss: 0.0007213834905996919
step: 40, loss: 0.0009466066840104759
step: 50, loss: 0.0006962153711356223
step: 60, loss: 0.0023831266444176435
step: 70, loss: 0.0011865360429510474
step: 80, loss: 0.004149154294282198
step: 90, loss: 0.004355488810688257
step: 100, loss: 0.002579648746177554
step: 110, loss: 0.0011457508662715554
step: 120, loss: 0.045681484043598175
step: 130, loss: 0.0009620178025215864
step: 140, loss: 0.006168849300593138
step: 150, loss: 0.0013989447616040707
step: 160, loss: 0.0009620531345717609
step: 170, loss: 0.0022137085907161236
step: 180, loss: 0.0012428261106833816
step: 190, loss: 0.005277146585285664
step: 200, loss: 0.0005677009467035532
step: 210, loss: 0.0005370258004404604
step: 220, loss: 0.0009009453933686018
step: 230, loss: 0.00038034829776734114
step: 240, loss: 0.009633179754018784
step: 250, loss: 0.0005888204323127866
step: 260, loss: 0.00040819469722919166
step: 270, loss: 0.0007753782556392252
step: 280, loss: 0.0003286390856374055
step: 290, loss: 0.00044992819312028587
step: 300, loss: 0.00035224496969021857
step: 310, loss: 0.001194618409499526
step: 320, loss: 0.00040702527621760964
step: 330, loss: 0.00029094977071508765
step: 340, loss: 0.0013210487086325884
step: 350, loss: 0.006255178712308407

```

```

step: 360, loss: 0.0003696104104164988
step: 370, loss: 0.0006422695587389171
step: 380, loss: 0.0013714596861973405
step: 390, loss: 0.00106335140299052
step: 400, loss: 0.0026668482460081577
step: 410, loss: 0.0015941762831062078
step: 420, loss: 0.0025477581657469273
step: 430, loss: 0.006508261896669865
step: 440, loss: 0.0011201683664694428
step: 450, loss: 0.0011104975128546357
step: 460, loss: 0.10057494044303894
step: 470, loss: 0.0012274312321096659
step: 480, loss: 0.0008248291560448706
step: 490, loss: 0.0015240737702697515
step: 500, loss: 0.0006384022999554873
step: 510, loss: 0.0006038458086550236
step: 520, loss: 0.0006338090752251446
step: 530, loss: 0.012425493448972702
/usr/local/lib/python3.7/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=0.9262180974477957, f1=0.9181395348837208,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007224069559015334
step: 10, loss: 0.01090015284717083
step: 20, loss: 0.014490167610347271
step: 30, loss: 0.00017372674483340234
step: 40, loss: 0.1413085013628006
step: 50, loss: 0.0007508599082939327
step: 60, loss: 0.000562961446121335
step: 70, loss: 0.0006626421236433089
step: 80, loss: 0.0005659498274326324
step: 90, loss: 0.0009538863669149578
step: 100, loss: 0.20820747315883636
step: 110, loss: 0.004952315241098404
step: 120, loss: 0.006543746683746576
step: 130, loss: 0.0014367976691573858
step: 140, loss: 0.0005375643377192318
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 150, loss: 0.0003414075472392142
step: 160, loss: 0.0016804822953417897
step: 170, loss: 0.00033698847983032465
step: 180, loss: 0.000721242802683264
step: 190, loss: 0.001100299647077918
step: 200, loss: 0.0004296144179534167

```

```

step: 210, loss: 0.0007861628546379507
step: 220, loss: 0.001300671137869358
step: 230, loss: 0.0006327331066131592
step: 240, loss: 0.0007039280026219785
step: 250, loss: 0.0005763796507380903
step: 260, loss: 0.0014594290405511856
step: 270, loss: 0.0015023835003376007
step: 280, loss: 0.0034452860709279776
step: 290, loss: 0.002592191332951188
step: 300, loss: 0.000997656607069075
step: 310, loss: 0.0010708183981478214
step: 320, loss: 0.00904066115617752
step: 330, loss: 0.0037881885655224323
step: 340, loss: 0.0008652397082187235
step: 350, loss: 0.0003342572890687734
step: 360, loss: 0.0008293531718663871
step: 370, loss: 0.0015646808315068483
step: 380, loss: 0.0010802499018609524
step: 390, loss: 0.0009582058992236853
step: 400, loss: 0.000721910095307976
step: 410, loss: 0.01830342598259449
step: 420, loss: 0.0020964930299669504
step: 430, loss: 0.0003664351534098387
step: 440, loss: 0.12830710411071777
step: 450, loss: 0.0032455851323902607
step: 460, loss: 0.0009437738917768002
step: 470, loss: 0.002670567948371172
step: 480, loss: 0.013421609066426754
step: 490, loss: 0.0018154517747461796
step: 500, loss: 0.0003407628391869366
step: 510, loss: 0.004466691054403782
step: 520, loss: 0.008841187693178654
step: 530, loss: 0.0004812315455637872
/usr/local/lib/python3.7/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=0.9296693060083837, f1=0.9140989729225022,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002481637638993561
step: 10, loss: 0.0004572120087686926
step: 20, loss: 0.0009201596258208156
step: 30, loss: 0.00020814039453398436
step: 40, loss: 0.05300331115722656
step: 50, loss: 0.0005102620925754309
step: 60, loss: 0.00020506835426203907

```

step: 70, loss: 0.030152752995491028
step: 80, loss: 0.0012379074469208717
step: 90, loss: 0.0003209621936548501
step: 100, loss: 0.00025843470939435065
step: 110, loss: 0.00021526761702261865
step: 120, loss: 0.0002377385535510257
step: 130, loss: 0.0003538618329912424
step: 140, loss: 0.002209064085036516
step: 150, loss: 0.0006673352909274399
step: 160, loss: 0.0013095615431666374
step: 170, loss: 0.0001429217663826421
step: 180, loss: 0.000288552837446332
step: 190, loss: 0.0003207878617104143
step: 200, loss: 0.00037301413249224424
step: 210, loss: 0.0006447884952649474
step: 220, loss: 0.000422877463279292
step: 230, loss: 0.0008195508853532374
step: 240, loss: 0.00048420947859995067
step: 250, loss: 0.015629548579454422
step: 260, loss: 0.2328745722770691
step: 270, loss: 0.0005358998314477503
step: 280, loss: 0.05685415863990784
step: 290, loss: 0.0003743474662769586
step: 300, loss: 0.0005362536758184433
step: 310, loss: 0.0009394206572324038
step: 320, loss: 0.0003934031701646745
step: 330, loss: 0.0006195915630087256
step: 340, loss: 0.0015737836947664618
step: 350, loss: 0.0005035026115365326
step: 360, loss: 0.18910248577594757
step: 370, loss: 0.0012766099534928799
step: 380, loss: 0.0008228283259086311
step: 390, loss: 0.0009912890382111073
step: 400, loss: 0.000844735826831311
step: 410, loss: 0.002197731053456664
step: 420, loss: 0.0007403908530250192
step: 430, loss: 0.0007283115992322564
step: 440, loss: 0.00027453788788989186
step: 450, loss: 0.0020819876808673143
step: 460, loss: 0.004039658233523369
step: 470, loss: 0.0008969917544163764
step: 480, loss: 0.0008910810574889183
step: 490, loss: 0.00035224881139583886
step: 500, loss: 0.0003796280943788588
step: 510, loss: 0.025644607841968536
step: 520, loss: 0.0003434997342992574
step: 530, loss: 0.0004189138126093894
/usr/local/lib/python3.7/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=0.9263351749539595, f1=0.911955514365153,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003234915202483535
step: 10, loss: 0.0014188281493261456
step: 20, loss: 0.0005624886252917349
step: 30, loss: 0.0006906106136739254
step: 40, loss: 0.0004502391966525465
step: 50, loss: 0.19030477106571198
step: 60, loss: 0.002352580428123474
step: 70, loss: 0.001303448691032827
step: 80, loss: 0.00111842539627105
step: 90, loss: 0.00036680669290944934
step: 100, loss: 0.0014112937496975064
step: 110, loss: 0.0006169176776893437
step: 120, loss: 0.0007355424459092319
step: 130, loss: 0.00043868570355698466
step: 140, loss: 0.0008254295098595321
step: 150, loss: 0.001227573724463582
step: 160, loss: 0.0004047782567795366
step: 170, loss: 0.001410845317877829
step: 180, loss: 0.00023575829982291907
step: 190, loss: 0.001405783579684794
step: 200, loss: 0.0006866236217319965
step: 210, loss: 0.0013634293572977185
step: 220, loss: 0.00048180672456510365
step: 230, loss: 0.00048452967894263566
step: 240, loss: 0.001040223753079772
step: 250, loss: 0.0018154600402340293
step: 260, loss: 0.00045328980195336044
step: 270, loss: 0.0006171968416310847
step: 280, loss: 0.0003933035477530211
step: 290, loss: 0.00017912450130097568
step: 300, loss: 0.0012911342782899737
step: 310, loss: 0.00022004585480317473
step: 320, loss: 0.0004568055737763643
step: 330, loss: 0.0006474683759734035
step: 340, loss: 0.00034438882721588016
step: 350, loss: 0.0003183943044859916
step: 360, loss: 0.0006817459361627698
step: 370, loss: 0.0017455037450417876
step: 380, loss: 0.04636595770716667
step: 390, loss: 0.00019484921358525753
step: 400, loss: 0.0012075472623109818
```

```
step: 410, loss: 0.00016909727128222585
step: 420, loss: 0.0005693468847312033
step: 430, loss: 0.0009186483221128583
step: 440, loss: 0.0009824368171393871
step: 450, loss: 0.00048574325046502054
step: 460, loss: 0.004200538620352745
step: 470, loss: 0.00041323748882859945
step: 480, loss: 0.0004830828111153096
step: 490, loss: 0.0005167472991161048
step: 500, loss: 0.000609254464507103
step: 510, loss: 0.0009029911016114056
step: 520, loss: 0.0002321530773770064
step: 530, loss: 0.00039333465974777937
/usr/local/lib/python3.7/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=0.9287377736376339, f1=0.9178789300797747,
best_f1=0.9133271202236719
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003389054036233574
step: 10, loss: 0.010688497684895992
step: 20, loss: 0.000496537599246949
step: 30, loss: 0.0018274645553901792
step: 40, loss: 0.0003695031045936048
step: 50, loss: 0.0008394865435548127
step: 60, loss: 0.0010318681597709656
step: 70, loss: 0.0005070038023404777
step: 80, loss: 0.00040199767681770027
step: 90, loss: 0.0005562249571084976
step: 100, loss: 0.00012262686504982412
step: 110, loss: 0.002515587490051985
step: 120, loss: 0.0002685635699890554
step: 130, loss: 0.0013392339460551739
step: 140, loss: 0.000940683763474226
step: 150, loss: 0.0004253521328791976
step: 160, loss: 0.010339401662349701
step: 170, loss: 0.002126851584762335
step: 180, loss: 0.0011412949534133077
step: 190, loss: 0.000584052933845669
step: 200, loss: 0.0001985540147870779
step: 210, loss: 0.0005245613865554333
step: 220, loss: 0.0006193115841597319
step: 230, loss: 0.0008285414660349488
step: 240, loss: 0.00024941086303442717
step: 250, loss: 0.0010076462058350444
step: 260, loss: 0.026229223236441612
```

```

step: 270, loss: 0.0005168567295186222
step: 280, loss: 0.00022731476929038763
step: 290, loss: 0.0004695513052865863
step: 300, loss: 0.0005406475975178182
step: 310, loss: 0.03237896412611008
step: 320, loss: 0.00035348202800378203
step: 330, loss: 0.000636004377156496
step: 340, loss: 0.0004586479626595974
step: 350, loss: 0.0014043146511539817
step: 360, loss: 0.009444328024983406
step: 370, loss: 0.0008139002602547407
step: 380, loss: 0.001232022768817842
step: 390, loss: 0.0005293911672197282
step: 400, loss: 0.01427729893475771
step: 410, loss: 0.0011535121593624353
step: 420, loss: 0.0005790971335954964
step: 430, loss: 0.00019264676666352898
step: 440, loss: 0.00115727458614856
step: 450, loss: 0.0035630532074719667
step: 460, loss: 0.002999075222760439
step: 470, loss: 0.0002702352940104902
step: 480, loss: 0.006885637994855642
step: 490, loss: 0.0012845188612118363
step: 500, loss: 0.00023966087610460818
step: 510, loss: 0.0002794433676172048
step: 520, loss: 0.00033958189305849373
step: 530, loss: 0.0002452280605211854
/usr/local/lib/python3.7/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=0.9301453352086263, f1=0.9139072847682119,
best_f1=0.9133271202236719
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 [00:22, 251.13it/s]
load_f1 = 0.9327231121281464
real_f1 = 0.9287696577243294
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 [00:20, 215.29it/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 496 kB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: 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: 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: 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: 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: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: 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: 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: 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: 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: 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: 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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)

```

```

Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in
/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=63d9ea4133620f36d532cdbcdf8e76c3bbbf55a15da784732fa9a5b879f163db
  Stored in directory: /tmp/pip-ephem-wheel-cache-
rj79mlb1/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

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

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

```

```
--lm roberta \  
--use_gpu \  
--fp16 \  
--checkpoint_path checkpoints/
```

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: 0.442573606967926

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=0.4333333333333333, f1=0.4186046511627907,
best_f1=0.4186046511627907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4974696934223175
/usr/local/lib/python3.7/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=0.37209302325581395, f1=0.3225806451612903,
best_f1=0.4186046511627907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4251985549926758
/usr/local/lib/python3.7/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=0.3333333333333333, f1=0.36363636363636365,
best_f1=0.4186046511627907
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2997533977031708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.46153846153846156, f1=0.4285714285714286,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3065313696861267
/usr/local/lib/python3.7/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=0.41791044776119407, f1=0.3846153846153846,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
```

```

step: 0, loss: 0.246802419424057
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.43999999999999995, f1=0.43902439024390244,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5252156257629395
/usr/local/lib/python3.7/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=0.4117647058823529, f1=0.4363636363636364,
best_f1=0.4285714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.456521600484848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5531914893617021, f1=0.5853658536585367,
best_f1=0.5853658536585367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38176339864730835
/usr/local/lib/python3.7/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=0.5531914893617021, f1=0.5909090909090909,
best_f1=0.5853658536585367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.49750906229019165
/usr/local/lib/python3.7/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=0.5263157894736842, f1=0.4736842105263159,
best_f1=0.5853658536585367
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.44134050607681274
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6046511627906977, f1=0.5641025641025641,
best_f1=0.5641025641025641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28847476840019226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6341463414634146, f1=0.631578947368421,
best_f1=0.631578947368421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3156852126121521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.6857142857142857,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18991759419441223
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.6857142857142857,
best_f1=0.6857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3327644169330597
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.6857142857142857,
best_f1=0.6857142857142857
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
91it [00:00, 122687.77it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.6086956521739131
real_f1 = 0.5283018867924528
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:20, 213.02it/s]

```

3.1.2 DBLP-ACM - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:


```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5868017673492432
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.5303964018821716
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.6087181568145752
step: 30, loss: 0.30104970932006836
step: 40, loss: 0.3203975558280945
step: 50, loss: 0.43472784757614136
step: 60, loss: 0.16723884642124176
step: 70, loss: 0.06282869726419449
step: 80, loss: 0.04541340097784996
step: 90, loss: 0.08021446317434311
step: 100, loss: 0.14491596817970276
step: 110, loss: 0.16777876019477844
step: 120, loss: 0.05940144509077072
step: 130, loss: 0.047851987183094025
step: 140, loss: 0.01721017062664032
step: 150, loss: 0.08284691721200943
step: 160, loss: 0.2171693742275238
step: 170, loss: 0.04499254375696182
step: 180, loss: 0.00943361222743988
step: 190, loss: 0.02620837092399597
step: 200, loss: 0.020276522263884544
step: 210, loss: 0.01957329921424389
step: 220, loss: 0.014484627172350883
step: 230, loss: 0.03532220423221588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9596510359869138, f1=0.9604395604395605,

```

```

best_f1=0.9604395604395605
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013290779665112495
step: 10, loss: 0.2455165535211563
step: 20, loss: 0.03412420675158501
step: 30, loss: 0.04547726362943649
step: 40, loss: 0.05412640422582626
step: 50, loss: 0.016161508858203888
step: 60, loss: 0.012289069592952728
step: 70, loss: 0.050241127610206604
step: 80, loss: 0.011585721746087074
step: 90, loss: 0.027680393308401108
step: 100, loss: 0.006637485697865486
step: 110, loss: 0.005179265979677439
step: 120, loss: 0.0224575437605381
step: 130, loss: 0.005043352022767067
step: 140, loss: 0.04216935113072395
step: 150, loss: 0.055152587592601776
step: 160, loss: 0.0018716754857450724
step: 170, loss: 0.0023696941789239645
step: 180, loss: 0.00797708984464407
step: 190, loss: 0.2595987617969513
step: 200, loss: 0.0065934802405536175
step: 210, loss: 0.037545058876276016
step: 220, loss: 0.0020313311833888292
step: 230, loss: 0.0027768106665462255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, f1=0.9730337078651685,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01397774275392294
step: 10, loss: 0.01180912833660841
step: 20, loss: 0.0008765783859416842
step: 30, loss: 0.0006051703239791095
step: 40, loss: 0.0052988058887422085
step: 50, loss: 0.03160085529088974
step: 60, loss: 0.00334958010353148
step: 70, loss: 0.0025895400904119015

```

```

step: 80, loss: 0.03609447926282883
step: 90, loss: 0.08423855155706406
step: 100, loss: 0.014950834214687347
step: 110, loss: 0.0013560457155108452
step: 120, loss: 0.0006955372518859804
step: 130, loss: 0.004031600896269083
step: 140, loss: 0.0034982645884156227
step: 150, loss: 0.11142630130052567
step: 160, loss: 0.006264129187911749
step: 170, loss: 0.001692923135124147
step: 180, loss: 0.012444669380784035
step: 190, loss: 0.01121970359236002
step: 200, loss: 0.021874431520700455
step: 210, loss: 0.07960058003664017
step: 220, loss: 0.061528462916612625
step: 230, loss: 0.010821610689163208
/usr/local/lib/python3.7/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=0.9574944071588367, f1=0.9491150442477876,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061930324882268906
step: 10, loss: 0.006982884369790554
step: 20, loss: 0.007096052169799805
step: 30, loss: 0.0009750334429554641
step: 40, loss: 0.1521649807691574
step: 50, loss: 0.005416335538029671
step: 60, loss: 0.00175286119338125
step: 70, loss: 0.011995486915111542
step: 80, loss: 0.000675040006171912
step: 90, loss: 0.0026299715973436832
step: 100, loss: 0.0036076055839657784
step: 110, loss: 0.0032138878013938665
step: 120, loss: 0.003285262268036604
step: 130, loss: 0.008400422520935535
step: 140, loss: 0.028325233608484268
step: 150, loss: 0.040680211037397385
step: 160, loss: 0.0005384706892073154
step: 170, loss: 0.0058127534575760365
step: 180, loss: 0.044764623045921326
step: 190, loss: 0.04684678837656975
step: 200, loss: 0.1949651539325714
step: 210, loss: 0.0018178779864683747
step: 220, loss: 0.01043738704174757
step: 230, loss: 0.040904466062784195

```

```

/usr/local/lib/python3.7/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=0.9697648376259798, f1=0.9830890642615557,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007497934275306761
step: 10, loss: 0.0008794284658506513
step: 20, loss: 0.0011738031171262264
step: 30, loss: 0.00272737885825336
step: 40, loss: 0.0007387110381387174
step: 50, loss: 0.014791597612202168
step: 60, loss: 0.0019666836597025394
step: 70, loss: 0.002929435111582279
step: 80, loss: 0.01844063214957714
step: 90, loss: 0.04095025733113289
step: 100, loss: 0.0006879913853481412
step: 110, loss: 0.008250445127487183
step: 120, loss: 0.10183583199977875
step: 130, loss: 0.0034721402917057276
step: 140, loss: 0.05197524651885033
step: 150, loss: 0.00802411139011383
step: 160, loss: 0.0018514409894123673
step: 170, loss: 0.002645130967721343
step: 180, loss: 0.00828650314360857
step: 190, loss: 0.005843111779540777
step: 200, loss: 0.0051979669369757175
step: 210, loss: 0.0012892084196209908
step: 220, loss: 0.0012615922605618834
step: 230, loss: 0.003048193408176303
/usr/local/lib/python3.7/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=0.9764837625979844, f1=0.9808773903262092,
best_f1=0.9730337078651685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007348362705670297
step: 10, loss: 0.002790065249428153
step: 20, loss: 0.00364795234054327
step: 30, loss: 0.0024013358633965254
step: 40, loss: 0.001189662260003388
step: 50, loss: 0.001305465935729444
step: 60, loss: 0.0036575093399733305
step: 70, loss: 0.006998412776738405

```

```

step: 80, loss: 0.12646058201789856
step: 90, loss: 0.01766761764883995
step: 100, loss: 0.0011661634780466557
step: 110, loss: 0.02380548045039177
step: 120, loss: 0.0004422999045345932
step: 130, loss: 0.000555702019482851
step: 140, loss: 0.0015774747589603066
step: 150, loss: 0.0003180886269547045
step: 160, loss: 0.0009619990014471114
step: 170, loss: 0.0017880311934277415
step: 180, loss: 0.0006329126772470772
step: 190, loss: 0.002822005655616522
step: 200, loss: 0.023332417011260986
step: 210, loss: 0.0026175417006015778
step: 220, loss: 0.004075496457517147
step: 230, loss: 0.0024504815228283405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.978865406006674, f1=0.9798206278026906,
best_f1=0.9798206278026906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002296363702043891
step: 10, loss: 0.0018591789994388819
step: 20, loss: 0.0007798050064593554
step: 30, loss: 0.0005642481846734881
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 40, loss: 0.00099333340767771
step: 50, loss: 0.0008932147175073624
step: 60, loss: 0.00045663898345083
step: 70, loss: 0.0012990310788154602
step: 80, loss: 0.00958824809640646
step: 90, loss: 0.00205918587744236
step: 100, loss: 0.0005520335980691016
step: 110, loss: 0.0008898924570530653
step: 120, loss: 0.0004636074008885771
step: 130, loss: 0.0006996559677645564
step: 140, loss: 0.0005792079027742147
step: 150, loss: 0.0016459410544484854
step: 160, loss: 0.0004831101978197694
step: 170, loss: 0.0022721595596522093
step: 180, loss: 0.0016419498715549707

```

```

step: 190, loss: 0.00035390237462706864
step: 200, loss: 0.03647252544760704
step: 210, loss: 0.013854135759174824
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 2048.0
step: 220, loss: 0.16118748486042023
step: 230, loss: 0.0013469394762068987
/usr/local/lib/python3.7/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=0.9670329670329669, f1=0.977728285077951,
best_f1=0.9798206278026906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009993293788284063
step: 10, loss: 0.024071257561445236
step: 20, loss: 0.0020055631175637245
step: 30, loss: 0.0013372896937653422
step: 40, loss: 0.0015180015470832586
step: 50, loss: 0.0005407326389104128
step: 60, loss: 0.0006784183206036687
step: 70, loss: 0.00027044740272685885
step: 80, loss: 0.006557963322848082
step: 90, loss: 0.00029504502890631557
step: 100, loss: 0.0010029622353613377
step: 110, loss: 0.004209799692034721
step: 120, loss: 0.015535845421254635
step: 130, loss: 0.0025836117565631866
step: 140, loss: 0.0007292920490726829
step: 150, loss: 0.03120441734790802
step: 160, loss: 0.004663487896323204
step: 170, loss: 0.014887184835970402
step: 180, loss: 0.004877225961536169
step: 190, loss: 0.004139437805861235
step: 200, loss: 0.0014018821530044079
step: 210, loss: 0.00389050948433578
step: 220, loss: 0.0019125493708997965
step: 230, loss: 0.0012927408097311854
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9810901001112348, f1=0.9798657718120806,
best_f1=0.9798657718120806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011342902667820454
step: 10, loss: 0.001331416191533208
step: 20, loss: 0.001285469625145197
step: 30, loss: 0.0012728988658636808
step: 40, loss: 0.0008129930938594043
step: 50, loss: 0.003211717354133725
step: 60, loss: 0.006032873410731554
step: 70, loss: 0.0595279298722744
step: 80, loss: 0.115447998046875
step: 90, loss: 0.005095178727060556
step: 100, loss: 0.0005259407334960997
step: 110, loss: 0.0020164574962109327
step: 120, loss: 0.03541059046983719
step: 130, loss: 0.018584605306386948
step: 140, loss: 0.0006872297963127494
step: 150, loss: 0.0008931467309594154
step: 160, loss: 0.0015811027260497212
step: 170, loss: 0.0007837883895263076
step: 180, loss: 0.0007285673054866493
step: 190, loss: 0.0004908624105155468
step: 200, loss: 0.06622958183288574
step: 210, loss: 0.04275340959429741
step: 220, loss: 0.002901152940467
step: 230, loss: 0.0012109605595469475
/usr/local/lib/python3.7/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=0.9774774774774775, f1=0.9831271091113611,
best_f1=0.9798657718120806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013079928467050195
step: 10, loss: 0.022257603704929352
step: 20, loss: 0.0019216835498809814
step: 30, loss: 0.00016439439787063748
step: 40, loss: 0.0021090111695230007
step: 50, loss: 0.000338671263307333
step: 60, loss: 0.0009934710105881095
step: 70, loss: 0.0870940238237381
step: 80, loss: 0.001984131056815386
step: 90, loss: 0.0007131509482860565
step: 100, loss: 0.0014120929408818483
step: 110, loss: 0.004426844883710146
step: 120, loss: 0.00022131310834083706
step: 130, loss: 0.00024363935517612845

```

```

step: 140, loss: 0.0003129101824015379
step: 150, loss: 0.0010688824113458395
step: 160, loss: 7.030742563074455e-05
step: 170, loss: 0.0001746952038956806
step: 180, loss: 0.0016722202999517322
step: 190, loss: 0.00027799050440080464
step: 200, loss: 0.00044227400212548673
step: 210, loss: 0.004355356562882662
step: 220, loss: 0.13140715658664703
step: 230, loss: 0.0003463605826254934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9787709497206705,
best_f1=0.9787709497206705
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019440145115368068
step: 10, loss: 0.0006165459053590894
step: 20, loss: 0.00033479565172456205
step: 30, loss: 0.00046682695392519236
step: 40, loss: 0.0001307904749410227
step: 50, loss: 0.00021073524840176105
step: 60, loss: 0.0034116506576538086
step: 70, loss: 0.006117125973105431
step: 80, loss: 0.00014765012019779533
step: 90, loss: 0.13744212687015533
step: 100, loss: 0.003067820332944393
step: 110, loss: 0.00040334544610232115
step: 120, loss: 0.0009524776833131909
step: 130, loss: 0.0007793134427629411
step: 140, loss: 0.006586987525224686
step: 150, loss: 0.0017480365931987762
step: 160, loss: 0.004907221533358097
step: 170, loss: 0.007366582751274109
step: 180, loss: 0.0008948651957325637
step: 190, loss: 0.00042942946311086416
step: 200, loss: 0.006564405281096697
step: 210, loss: 0.0005058203241787851
step: 220, loss: 0.0024796498473733664
step: 230, loss: 0.00022283545695245266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841628959276018, f1=0.9841628959276018,
best_f1=0.9841628959276018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000315613899147138
step: 10, loss: 0.00011087710299761966
step: 20, loss: 0.009711544029414654
step: 30, loss: 0.0743827074766159
step: 40, loss: 0.0005130994832143188
step: 50, loss: 0.002215523738414049
step: 60, loss: 0.00021712869056500494
step: 70, loss: 0.00014451405149884522
step: 80, loss: 7.258655386976898e-05
step: 90, loss: 0.001083848183043301
step: 100, loss: 0.00012310975580476224
step: 110, loss: 0.0001745995250530541
step: 120, loss: 0.0003206938854418695
step: 130, loss: 0.0006946310750208795
step: 140, loss: 0.010913748294115067
step: 150, loss: 0.00029247987549751997
step: 160, loss: 0.00016837204748298973
step: 170, loss: 0.00048221476026810706
step: 180, loss: 0.0009256932535208762
step: 190, loss: 0.004802991636097431
step: 200, loss: 0.00030947092454880476
step: 210, loss: 0.0003006780752912164
step: 220, loss: 0.000535497150849551
step: 230, loss: 0.00017282857152167708
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854096520763187, f1=0.9853768278965129,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029317886219359934
step: 10, loss: 8.628959767520428e-05

```

```

step: 20, loss: 0.00029415020253509283
step: 30, loss: 0.002576008439064026
step: 40, loss: 0.00021478786948136985
step: 50, loss: 0.00017038837540894747
step: 60, loss: 0.0001583555422257632
step: 70, loss: 0.00017058629600796849
step: 80, loss: 6.435388786485419e-05
step: 90, loss: 0.000292244425509125
step: 100, loss: 0.0005811178707517684
step: 110, loss: 0.00012808249448426068
step: 120, loss: 0.0003363295109011233
step: 130, loss: 0.00017524519353173673
step: 140, loss: 8.094128861557692e-05
step: 150, loss: 0.0006036101258359849
step: 160, loss: 0.007032713852822781
step: 170, loss: 0.0016666986048221588
step: 180, loss: 0.0026415165048092604
step: 190, loss: 0.00033414672361686826
step: 200, loss: 3.616835965658538e-05
step: 210, loss: 0.002853778889402747
step: 220, loss: 0.00011326429375912994
step: 230, loss: 0.00014629379438702017
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.984304932735426,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012389026232995093
step: 10, loss: 9.055442205863073e-05
step: 20, loss: 0.00011456471111159772
step: 30, loss: 0.00018645681848283857
step: 40, loss: 0.00012778873497154564
step: 50, loss: 0.00014306310913525522
step: 60, loss: 0.0001449325936846435
step: 70, loss: 8.630291267763823e-05
step: 80, loss: 8.72675227583386e-05
step: 90, loss: 0.0002272811980219558
step: 100, loss: 0.00023876252816990018
step: 110, loss: 0.00020595832029357553
step: 120, loss: 2.7643076464300975e-05
step: 130, loss: 0.00042614858830347657
step: 140, loss: 6.530038808705285e-05
step: 150, loss: 0.001122455345466733
step: 160, loss: 0.0002476402441971004
step: 170, loss: 0.00014146386820357293

```

```

step: 180, loss: 8.652765245642513e-05
step: 190, loss: 0.0012747830478474498
step: 200, loss: 0.0001506501721451059
step: 210, loss: 0.00022758122941013426
step: 220, loss: 0.00030278830672614276
step: 230, loss: 0.00028587071574293077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9854423292273236, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029267376521602273
step: 10, loss: 0.00015054189134389162
step: 20, loss: 0.0018352713668718934
step: 30, loss: 0.0009678143542259932
step: 40, loss: 4.235597953083925e-05
step: 50, loss: 0.00011033355258405209
step: 60, loss: 0.0432741604745388
step: 70, loss: 0.00010603765258565545
step: 80, loss: 0.00017636535631027073
step: 90, loss: 9.054451220436022e-05
step: 100, loss: 3.599856063374318e-05
step: 110, loss: 0.00024247141845989972
step: 120, loss: 0.00036779852234758437
step: 130, loss: 0.0001753930700942874
step: 140, loss: 0.006652530282735825
step: 150, loss: 0.00011618502321653068
step: 160, loss: 0.00015650827845092863
step: 170, loss: 8.913016790756956e-05
step: 180, loss: 0.00015852152137085795
step: 190, loss: 7.287576590897515e-05
step: 200, loss: 0.00014900797395966947
step: 210, loss: 0.0003649965801741928
step: 220, loss: 0.00033592397812753916
step: 230, loss: 0.00015512193203903735
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9854096520763187,
best_f1=0.9854096520763187
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
2473it [00:11, 214.28it/s]
load_f1 = 0.9832402234636871
real_f1 = 0.9821029082774049
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:20, 214.32it/s]
```

3.1.3 DBLP-GoogleScholar - Running the matcher

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

```

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

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6652434468269348
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4723648726940155
step: 20, loss: 0.3251165449619293
step: 30, loss: 0.44075268507003784
step: 40, loss: 0.4648902416229248
step: 50, loss: 0.6543816328048706
step: 60, loss: 0.3684346079826355
step: 70, loss: 0.4345562160015106
step: 80, loss: 0.4679276943206787
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 90, loss: 0.3820081651210785
step: 100, loss: 0.3790214955806732
step: 110, loss: 0.41244378685951233
step: 120, loss: 0.28037771582603455
step: 130, loss: 0.2622748911380768
step: 140, loss: 0.21284779906272888
step: 150, loss: 0.12382815778255463
step: 160, loss: 0.3034604489803314
step: 170, loss: 0.14161278307437897
step: 180, loss: 0.19843876361846924
step: 190, loss: 0.2014438658952713
step: 200, loss: 0.16640833020210266
step: 210, loss: 0.033055536448955536
step: 220, loss: 0.18455569446086884
step: 230, loss: 0.23359642922878265
step: 240, loss: 0.0844600647687912
step: 250, loss: 0.11345154047012329
step: 260, loss: 0.3098145127296448
step: 270, loss: 0.2867448627948761
step: 280, loss: 0.06341332942247391
step: 290, loss: 0.1375025510787964
step: 300, loss: 0.2712412178516388
step: 310, loss: 0.29679039120674133
step: 320, loss: 0.14931756258010864
step: 330, loss: 0.14169837534427643
step: 340, loss: 0.31841593980789185
step: 350, loss: 0.23580101132392883
step: 360, loss: 0.04506562277674675
step: 370, loss: 0.021174848079681396
step: 380, loss: 0.2879686653614044
step: 390, loss: 0.05099291726946831
step: 400, loss: 0.08034560084342957

```

```

step: 410, loss: 0.2512272000312805
step: 420, loss: 0.08409107476472855
step: 430, loss: 0.03145221620798111
step: 440, loss: 0.056381627917289734
step: 450, loss: 0.11576838791370392
step: 460, loss: 0.06704898178577423
step: 470, loss: 0.059442803263664246
step: 480, loss: 0.2177848368883133
step: 490, loss: 0.27024468779563904
step: 500, loss: 0.0638984739780426
step: 510, loss: 0.03490550071001053
step: 520, loss: 0.16892385482788086
step: 530, loss: 0.028710976243019104
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.906163753449862, f1=0.9023941068139963,
best_f1=0.9023941068139963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13568924367427826
step: 10, loss: 0.11631272733211517
step: 20, loss: 0.07974106818437576
step: 30, loss: 0.20445384085178375
step: 40, loss: 0.05096627399325371
step: 50, loss: 0.08128396421670914
step: 60, loss: 0.14563001692295074
step: 70, loss: 0.086742103099823
step: 80, loss: 0.1225394532084465
step: 90, loss: 0.05971783772110939
step: 100, loss: 0.22303406894207
step: 110, loss: 0.04270852729678154
step: 120, loss: 0.13825125992298126
step: 130, loss: 0.012318521738052368
step: 140, loss: 0.09326015412807465
step: 150, loss: 0.047542303800582886
step: 160, loss: 0.04348450154066086
step: 170, loss: 0.1250791847705841
step: 180, loss: 0.02680068276822567
step: 190, loss: 0.04614294692873955
step: 200, loss: 0.27687788009643555
step: 210, loss: 0.0353095717728138
step: 220, loss: 0.002431064611300826

```

```

step: 230, loss: 0.08076395094394684
step: 240, loss: 0.061559613794088364
step: 250, loss: 0.06734943389892578
step: 260, loss: 0.17093220353126526
step: 270, loss: 0.02478877268731594
step: 280, loss: 0.07643566280603409
step: 290, loss: 0.056656207889318466
step: 300, loss: 0.13572561740875244
step: 310, loss: 0.1088688001036644
step: 320, loss: 0.17212608456611633
step: 330, loss: 0.10441456735134125
step: 340, loss: 0.20957203209400177
step: 350, loss: 0.041226014494895935
step: 360, loss: 0.061398621648550034
step: 370, loss: 0.005252487026154995
step: 380, loss: 0.19427160918712616
step: 390, loss: 0.04802774265408516
step: 400, loss: 0.04003090411424637
step: 410, loss: 0.061395008116960526
step: 420, loss: 0.038138967007398605
step: 430, loss: 0.16709594428539276
step: 440, loss: 0.01975214295089245
step: 450, loss: 0.09016364067792892
step: 460, loss: 0.13600437343120575
step: 470, loss: 0.1453266441822052
step: 480, loss: 0.03592713922262192
step: 490, loss: 0.06781899929046631
step: 500, loss: 0.005575215443968773
step: 510, loss: 0.04542222619056702
step: 520, loss: 0.22093677520751953
step: 530, loss: 0.049904439598321915
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9219924812030075, f1=0.9171736078614879,
best_f1=0.9171736078614879
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11910157650709152
step: 10, loss: 0.0339091420173645
step: 20, loss: 0.01439501903951168
step: 30, loss: 0.07318463921546936
step: 40, loss: 0.09564206004142761

```


step: 50, loss: 0.03787769749760628
step: 60, loss: 0.13253434002399445
step: 70, loss: 0.020344141870737076
step: 80, loss: 0.07863129675388336
step: 90, loss: 0.01702469028532505
step: 100, loss: 0.14695514738559723
step: 110, loss: 0.03868431970477104
step: 120, loss: 0.06557601690292358
step: 130, loss: 0.07108256965875626
step: 140, loss: 0.003160880645737052
step: 150, loss: 0.04167594015598297
step: 160, loss: 0.0722230076789856
step: 170, loss: 0.03721383586525917
step: 180, loss: 0.033764760941267014
step: 190, loss: 0.04642583429813385
step: 200, loss: 0.060700833797454834
step: 210, loss: 0.057683784514665604
step: 220, loss: 0.04997160658240318
step: 230, loss: 0.05163269862532616
step: 240, loss: 0.055201176553964615
step: 250, loss: 0.18741309642791748
step: 260, loss: 0.036150284111499786
step: 270, loss: 0.009572066366672516
step: 280, loss: 0.00585085479542613
step: 290, loss: 0.021002132445573807
step: 300, loss: 0.23198045790195465
step: 310, loss: 0.08449673652648926
step: 320, loss: 0.026759102940559387
step: 330, loss: 0.022869639098644257
step: 340, loss: 0.03619345650076866
step: 350, loss: 0.1056692972779274
step: 360, loss: 0.0160038024187088
step: 370, loss: 0.16831931471824646
step: 380, loss: 0.10534704476594925
step: 390, loss: 0.032480206340551376
step: 400, loss: 0.07414699345827103
step: 410, loss: 0.023486468940973282
step: 420, loss: 0.024154437705874443
step: 430, loss: 0.09122572094202042
step: 440, loss: 0.16772383451461792
step: 450, loss: 0.05807707831263542
step: 460, loss: 0.08154414594173431
step: 470, loss: 0.011216538026928902
step: 480, loss: 0.11362702399492264
step: 490, loss: 0.029095645993947983
step: 500, loss: 0.020356081426143646
step: 510, loss: 0.06506579369306564
step: 520, loss: 0.05716366693377495

```

step: 530, loss: 0.02398582361638546
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9221501390176089, f1=0.9044117647058824,
best_f1=0.9044117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06410125643014908
step: 10, loss: 0.006770572159439325
step: 20, loss: 0.06706070154905319
step: 30, loss: 0.007599192205816507
step: 40, loss: 0.0045409733429551125
step: 50, loss: 0.010142285376787186
step: 60, loss: 0.00797477550804615
step: 70, loss: 0.08550787717103958
step: 80, loss: 0.030547192320227623
step: 90, loss: 0.09708571434020996
step: 100, loss: 0.001549137057736516
step: 110, loss: 0.04689177870750427
step: 120, loss: 0.004674754571169615
step: 130, loss: 0.031707219779491425
step: 140, loss: 0.08566943556070328
step: 150, loss: 0.11685473471879959
step: 160, loss: 0.029976096004247665
step: 170, loss: 0.01684706285595894
step: 180, loss: 0.05042900890111923
step: 190, loss: 0.0907582938671112
step: 200, loss: 0.044201262295246124
step: 210, loss: 0.014488345943391323
step: 220, loss: 0.05340532585978508
step: 230, loss: 0.00973384641110897
step: 240, loss: 0.007633610628545284
step: 250, loss: 0.12381824851036072
step: 260, loss: 0.02460431307554245
step: 270, loss: 0.046443697065114975
step: 280, loss: 0.02788027748465538
step: 290, loss: 0.010303042829036713
step: 300, loss: 0.008315439336001873
step: 310, loss: 0.002794332103803754
step: 320, loss: 0.20799019932746887
step: 330, loss: 0.014779730699956417
step: 340, loss: 0.010596890933811665

```

```

step: 350, loss: 0.07892747223377228
step: 360, loss: 0.0655096173286438
step: 370, loss: 0.013322332873940468
step: 380, loss: 0.006943321321159601
step: 390, loss: 0.007733731996268034
step: 400, loss: 0.03835929557681084
step: 410, loss: 0.13153104484081268
step: 420, loss: 0.0045766946859657764
step: 430, loss: 0.02405104972422123
step: 440, loss: 0.015138121321797371
step: 450, loss: 0.07332035154104233
step: 460, loss: 0.12692587077617645
step: 470, loss: 0.003769254544749856
step: 480, loss: 0.0035911768209189177
step: 490, loss: 0.10004901885986328
step: 500, loss: 0.11823298037052155
step: 510, loss: 0.02427057735621929
step: 520, loss: 0.004001827910542488
step: 530, loss: 0.03770942986011505
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9293577981651376, f1=0.9250457038391224,
best_f1=0.9250457038391224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003038073657080531
step: 10, loss: 0.07143709063529968
step: 20, loss: 0.025701604783535004
step: 30, loss: 0.0928143560886383
step: 40, loss: 0.009304207749664783
step: 50, loss: 0.04476756602525711
step: 60, loss: 0.10114248842000961
step: 70, loss: 0.017512034624814987
step: 80, loss: 0.0019089921843260527
step: 90, loss: 0.0024996509309858084
step: 100, loss: 0.08384035527706146
step: 110, loss: 0.0026897662319242954
step: 120, loss: 0.17986588180065155
step: 130, loss: 0.024313444271683693
step: 140, loss: 0.05985192954540253
step: 150, loss: 0.018312308937311172
step: 160, loss: 0.07584203034639359

```

```

step: 170, loss: 0.04349120706319809
step: 180, loss: 0.009843642823398113
step: 190, loss: 0.009072426706552505
step: 200, loss: 0.003558869007974863
step: 210, loss: 0.025682728737592697
step: 220, loss: 0.036200232803821564
step: 230, loss: 0.04872851446270943
step: 240, loss: 0.00720701040700078
step: 250, loss: 0.13156980276107788
step: 260, loss: 0.002026609843596816
step: 270, loss: 0.009641983546316624
step: 280, loss: 0.0158229973167181
step: 290, loss: 0.0449296273291111
step: 300, loss: 0.15182174742221832
step: 310, loss: 0.01315772533416748
step: 320, loss: 0.026125309988856316
step: 330, loss: 0.0013779832515865564
step: 340, loss: 0.0022629695013165474
step: 350, loss: 0.1163230836391449
step: 360, loss: 0.010776743292808533
step: 370, loss: 0.0207385066896677
step: 380, loss: 0.00191060162615031
step: 390, loss: 0.10117201507091522
step: 400, loss: 0.02471182309091091
step: 410, loss: 0.2065262794494629
step: 420, loss: 0.04735656827688217
step: 430, loss: 0.06014274060726166
step: 440, loss: 0.0027223071083426476
step: 450, loss: 0.011693427339196205
step: 460, loss: 0.0264419075101614
step: 470, loss: 0.00409238925203681
step: 480, loss: 0.011354085057973862
step: 490, loss: 0.042954180389642715
step: 500, loss: 0.012186449021100998
step: 510, loss: 0.0024094947148114443
step: 520, loss: 0.21046127378940582
step: 530, loss: 0.005172154866158962
/usr/local/lib/python3.7/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=0.9291412482402628, f1=0.9255813953488372,
best_f1=0.9250457038391224
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01074245199561119
step: 10, loss: 0.0025607554707676172
step: 20, loss: 0.006844340357929468

```

step: 30, loss: 0.001063216826878488
step: 40, loss: 0.004278147593140602
step: 50, loss: 0.001229635439813137
step: 60, loss: 0.01995936594903469
step: 70, loss: 0.0008386190747842193
step: 80, loss: 0.010272343643009663
step: 90, loss: 0.001753800199367106
step: 100, loss: 0.21992714703083038
step: 110, loss: 0.006425416562706232
step: 120, loss: 0.01178034022450447
step: 130, loss: 0.005073246080428362
step: 140, loss: 0.0026943490374833345
step: 150, loss: 0.0009849839843809605
step: 160, loss: 0.024984147399663925
step: 170, loss: 0.015144598670303822
step: 180, loss: 0.0016348648350685835
step: 190, loss: 0.005041792057454586
step: 200, loss: 0.04822163283824921
step: 210, loss: 0.00806631799787283
step: 220, loss: 0.002214849693700671
step: 230, loss: 0.01737535372376442
step: 240, loss: 0.005165571346879005
step: 250, loss: 0.12520739436149597
step: 260, loss: 0.005955324042588472
step: 270, loss: 0.0010787987848743796
step: 280, loss: 0.002759713213890791
step: 290, loss: 0.02851695753633976
step: 300, loss: 0.003748063463717699
step: 310, loss: 0.05909484624862671
step: 320, loss: 0.0004737502313219011
step: 330, loss: 0.020778784528374672
step: 340, loss: 0.0028225164860486984
step: 350, loss: 0.0029675087425857782
step: 360, loss: 0.004942090250551701
step: 370, loss: 0.003190811024978757
step: 380, loss: 0.0028619710355997086
step: 390, loss: 0.015389896929264069
step: 400, loss: 0.02219116874039173
step: 410, loss: 0.0037799563724547625
step: 420, loss: 0.033099498599767685
step: 430, loss: 0.056777093559503555
step: 440, loss: 0.00031104523804970086
step: 450, loss: 0.12157014012336731
step: 460, loss: 0.0023472988978028297
step: 470, loss: 0.009580710902810097
step: 480, loss: 0.005318242125213146
step: 490, loss: 0.1491769254207611
step: 500, loss: 0.010693352669477463

```

step: 510, loss: 0.24714191257953644
step: 520, loss: 0.0012973365373909473
step: 530, loss: 0.02101433277130127
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.936249418334109, f1=0.926012098650535,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013532016426324844
step: 10, loss: 0.008267749100923538
step: 20, loss: 0.00472627766430378
step: 30, loss: 0.009186208248138428
step: 40, loss: 0.03653456270694733
step: 50, loss: 0.0037722066044807434
step: 60, loss: 0.0011939076939597726
step: 70, loss: 0.03899817541241646
step: 80, loss: 0.0007500685169361532
step: 90, loss: 0.000448198348749429
step: 100, loss: 0.013616614043712616
step: 110, loss: 0.0005590365617536008
step: 120, loss: 0.0016974236350506544
step: 130, loss: 0.0005387827986851335
step: 140, loss: 0.011455214582383633
step: 150, loss: 0.00021555197599809617
step: 160, loss: 0.0005550140049308538
step: 170, loss: 0.0012459468562155962
step: 180, loss: 0.001834275433793664
step: 190, loss: 0.006786706857383251
step: 200, loss: 0.0008856943459250033
step: 210, loss: 0.05106797814369202
step: 220, loss: 0.022790241986513138
step: 230, loss: 0.01267692819237709
step: 240, loss: 0.033734843134880066
step: 250, loss: 0.018189413473010063
step: 260, loss: 0.0008867944125086069
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 270, loss: 0.0012373790377750993
step: 280, loss: 0.16415612399578094
step: 290, loss: 0.05092954263091087
step: 300, loss: 0.0015770135214552283
step: 310, loss: 0.004822412505745888

```

```

step: 320, loss: 0.004654985386878252
step: 330, loss: 0.13499537110328674
step: 340, loss: 0.019856184720993042
step: 350, loss: 0.012394482269883156
step: 360, loss: 0.0007632781635038555
step: 370, loss: 0.004821042064577341
step: 380, loss: 0.0021114007104188204
step: 390, loss: 0.007216375321149826
step: 400, loss: 0.09364328533411026
step: 410, loss: 0.0003110429388470948
step: 420, loss: 0.031645942479372025
step: 430, loss: 0.0049869897775352
step: 440, loss: 0.007007793057709932
step: 450, loss: 0.03342703357338905
step: 460, loss: 0.008003365248441696
step: 470, loss: 0.13267140090465546
step: 480, loss: 0.0025950309354811907
step: 490, loss: 0.0012425676686689258
step: 500, loss: 0.008805044926702976
step: 510, loss: 0.0029744955245405436
step: 520, loss: 0.004792392253875732
step: 530, loss: 0.002663536462932825
/usr/local/lib/python3.7/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=0.9341317365269461, f1=0.9238578680203047,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10251633822917938
step: 10, loss: 0.0018987864023074508
step: 20, loss: 0.0009049175423569977
step: 30, loss: 0.0015215068124234676
step: 40, loss: 0.00041211259667761624
step: 50, loss: 0.010256236419081688
step: 60, loss: 0.0007983961259014904
step: 70, loss: 0.0007842948543839157
step: 80, loss: 0.08442802727222443
step: 90, loss: 0.00243221502751112
step: 100, loss: 0.020285731181502342
step: 110, loss: 0.0003554256691131741
step: 120, loss: 0.003471873700618744
step: 130, loss: 0.0003454037068877369
step: 140, loss: 0.0037050163373351097
step: 150, loss: 0.0005379949579946697
step: 160, loss: 0.012787909246981144
step: 170, loss: 0.08206281065940857

```

```

step: 180, loss: 0.008695210330188274
step: 190, loss: 0.08162462711334229
step: 200, loss: 0.00307994126342237
step: 210, loss: 0.04374394565820694
step: 220, loss: 0.0044361078180372715
step: 230, loss: 0.0009545075590722263
step: 240, loss: 0.024090319871902466
step: 250, loss: 0.00027645655791275203
step: 260, loss: 0.00048602320021018386
step: 270, loss: 0.11893031001091003
step: 280, loss: 0.005001573823392391
step: 290, loss: 0.05667665973305702
step: 300, loss: 0.0006684773834422231
step: 310, loss: 0.001207932480610907
step: 320, loss: 0.001677391934208572
step: 330, loss: 0.0010638624662533402
step: 340, loss: 0.007252648007124662
step: 350, loss: 0.0009446752374060452
step: 360, loss: 0.0021107823122292757
step: 370, loss: 0.3379226326942444
step: 380, loss: 0.0019549785647541285
step: 390, loss: 0.0038065172266215086
step: 400, loss: 0.001132859499193728
step: 410, loss: 0.004183936398476362
step: 420, loss: 0.00784701481461525
step: 430, loss: 0.0017089064931496978
step: 440, loss: 0.005389206111431122
step: 450, loss: 0.0027500446885824203
step: 460, loss: 0.00708260340616107
step: 470, loss: 0.02156652882695198
step: 480, loss: 0.00907844491302967
step: 490, loss: 0.0017340043559670448
step: 500, loss: 0.005656599532812834
step: 510, loss: 0.05272187665104866
step: 520, loss: 0.0017019484657794237
step: 530, loss: 0.0079165855422616
/usr/local/lib/python3.7/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=0.9311141932501156, f1=0.9301895515487749,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.360196418128908e-05
step: 10, loss: 0.001448913593776524
step: 20, loss: 0.01866241730749607
step: 30, loss: 0.05329504236578941

```


step: 40, loss: 0.0005037139053456485
step: 50, loss: 0.004001085180789232
step: 60, loss: 0.00048413011245429516
step: 70, loss: 0.00021912208467256278
step: 80, loss: 0.03998805955052376
step: 90, loss: 0.0006928802467882633
step: 100, loss: 0.0070956978015601635
step: 110, loss: 0.04452894255518913
step: 120, loss: 0.0003042797325178981
step: 130, loss: 0.000860055792145431
step: 140, loss: 0.0006911581149324775
step: 150, loss: 0.05041806399822235
step: 160, loss: 0.03286856412887573
step: 170, loss: 0.005993454717099667
step: 180, loss: 0.033293452113866806
step: 190, loss: 0.00012009111378574744
step: 200, loss: 0.0003186861576978117
step: 210, loss: 0.21579857170581818
step: 220, loss: 0.004441088531166315
step: 230, loss: 0.000939554360229522
step: 240, loss: 0.0005486804293468595
step: 250, loss: 0.012025812640786171
step: 260, loss: 0.00033107836497947574
step: 270, loss: 0.0007033527945168316
step: 280, loss: 0.0019039945909753442
step: 290, loss: 0.00032945440034382045
step: 300, loss: 0.0008935061050578952
step: 310, loss: 0.036860208958387375
step: 320, loss: 0.0009992364794015884
step: 330, loss: 0.0036589426454156637
step: 340, loss: 0.0029846022371202707
step: 350, loss: 0.01778321899473667
step: 360, loss: 0.015044387429952621
step: 370, loss: 0.004939831793308258
step: 380, loss: 0.0005113205406814814
step: 390, loss: 0.000595874444115907
step: 400, loss: 0.0021195935551077127
step: 410, loss: 0.00040412115049548447
step: 420, loss: 0.0006268964498303831
step: 430, loss: 0.0011675709392875433
step: 440, loss: 0.00017123576253652573
step: 450, loss: 0.13574114441871643
step: 460, loss: 0.04315014183521271
step: 470, loss: 0.012213274836540222
step: 480, loss: 0.000214169587707147
step: 490, loss: 0.012577357701957226
step: 500, loss: 0.0010820586467161775
step: 510, loss: 0.013289852067828178

```

step: 520, loss: 0.002276989398524165
step: 530, loss: 0.13666768372058868
/usr/local/lib/python3.7/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=0.9332706766917293, f1=0.9296435272045028,
best_f1=0.926012098650535
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002233044244349003
step: 10, loss: 0.0006585845840163529
step: 20, loss: 0.0014817247865721583
step: 30, loss: 0.003936532884836197
step: 40, loss: 0.0002843996917363256
step: 50, loss: 0.0004174987552687526
step: 60, loss: 0.00608689384534955
step: 70, loss: 0.00032161775743588805
step: 80, loss: 0.0009380432311445475
step: 90, loss: 0.002517521847039461
step: 100, loss: 0.0008449433371424675
step: 110, loss: 0.0009525961359031498
step: 120, loss: 8.161891310010105e-05
step: 130, loss: 0.044820960611104965
step: 140, loss: 0.0007165326969698071
step: 150, loss: 0.03168732672929764
step: 160, loss: 0.006255232263356447
step: 170, loss: 0.0025728829205036163
step: 180, loss: 0.0013315402902662754
step: 190, loss: 9.25147978705354e-05
step: 200, loss: 0.0005692057893611491
step: 210, loss: 0.00011691252439050004
step: 220, loss: 0.00027528367354534566
step: 230, loss: 0.0001556008355692029
step: 240, loss: 0.0004044604138471186
step: 250, loss: 0.006842257920652628
step: 260, loss: 0.003219913924112916
step: 270, loss: 0.0009478587890043855
step: 280, loss: 0.01679692044854164
step: 290, loss: 0.00993403885513544
step: 300, loss: 0.07509253174066544
step: 310, loss: 0.16540805995464325
step: 320, loss: 0.0012295753695070744
step: 330, loss: 0.07475519180297852
step: 340, loss: 0.0025294432416558266
step: 350, loss: 0.006960299331694841
step: 360, loss: 0.0004451801360119134
step: 370, loss: 0.008568075485527515

```

```

step: 380, loss: 0.000937555218115449
step: 390, loss: 0.00024157523876056075
step: 400, loss: 0.0007827997324056923
step: 410, loss: 0.01085786521434784
step: 420, loss: 0.0002636596036609262
step: 430, loss: 0.0031182426027953625
step: 440, loss: 0.00023377645993605256
step: 450, loss: 0.059627484530210495
step: 460, loss: 0.00015435484237968922
step: 470, loss: 0.0002512590435799211
step: 480, loss: 0.0017431197920814157
step: 490, loss: 0.007740362547338009
step: 500, loss: 0.05410348251461983
step: 510, loss: 0.00012732816685456783
step: 520, loss: 0.048080217093229294
step: 530, loss: 0.003105292795225978
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9374143444495202, f1=0.9265845873233014,
best_f1=0.9265845873233014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014573815860785544
step: 10, loss: 0.00018986838404089212
step: 20, loss: 0.0005102744325995445
step: 30, loss: 0.0011212604586035013
step: 40, loss: 9.526668873149902e-05
step: 50, loss: 0.0001429775875294581
step: 60, loss: 0.0002579215506557375
step: 70, loss: 0.0007777144201099873
step: 80, loss: 0.0006349141476675868
step: 90, loss: 0.002176721580326557
step: 100, loss: 0.00012796904775314033
step: 110, loss: 0.0001221046841237694
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.0018132012337446213
step: 130, loss: 0.00020621228031814098
step: 140, loss: 0.00014624203322455287
step: 150, loss: 0.1036423072218895
step: 160, loss: 0.03203440085053444
step: 170, loss: 0.005892599932849407
step: 180, loss: 7.146002462832257e-05

```

```
step: 190, loss: 0.0035539132077246904
step: 200, loss: 0.00010932824807241559
step: 210, loss: 0.0016706384485587478
step: 220, loss: 0.00045010101166553795
step: 230, loss: 0.0018901810981333256
step: 240, loss: 0.007278529927134514
step: 250, loss: 0.0022736890241503716
step: 260, loss: 0.0003669720608741045
step: 270, loss: 0.0010146201821044087
step: 280, loss: 0.0007348342915065587
step: 290, loss: 6.830389611423016e-05
step: 300, loss: 6.655407923972234e-05
step: 310, loss: 0.00045117721310816705
step: 320, loss: 0.0001350288075627759
step: 330, loss: 0.0003500218444969505
step: 340, loss: 0.006718043703585863
step: 350, loss: 9.724259143695235e-05
step: 360, loss: 0.0002671251422725618
step: 370, loss: 0.00024880186538212
step: 380, loss: 0.023838786408305168
step: 390, loss: 0.00027955856057815254
step: 400, loss: 0.000160664101713337
step: 410, loss: 0.00011697543232003227
step: 420, loss: 0.0038683153688907623
step: 430, loss: 0.0002030300092883408
step: 440, loss: 0.00033180328318849206
step: 450, loss: 0.0012222277000546455
step: 460, loss: 0.0004109964065719396
step: 470, loss: 0.0002985477331094444
step: 480, loss: 0.0007124762050807476
step: 490, loss: 0.00020493954070843756
step: 500, loss: 0.0004403766943141818
step: 510, loss: 0.0013120429357513785
step: 520, loss: 0.0013671047054231167
step: 530, loss: 0.005624587647616863
/usr/local/lib/python3.7/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=0.9317865429234339, f1=0.9275092936802973,
best_f1=0.9265845873233014
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023770674306433648
step: 10, loss: 0.002088862005621195
step: 20, loss: 0.0011597013799473643
step: 30, loss: 0.0013171113096177578
step: 40, loss: 0.0030573552940040827
```

step: 50, loss: 0.00027022062567994
step: 60, loss: 0.00028604999533854425
step: 70, loss: 0.004868556745350361
step: 80, loss: 0.21745851635932922
step: 90, loss: 0.009377523325383663
step: 100, loss: 0.003962954506278038
step: 110, loss: 0.0006088009686209261
step: 120, loss: 0.0007992405444383621
step: 130, loss: 0.0008001121459528804
step: 140, loss: 0.0003970330872107297
step: 150, loss: 0.02950785495340824
step: 160, loss: 0.004720507655292749
step: 170, loss: 0.000967629486694932
step: 180, loss: 0.0007756820996291935
step: 190, loss: 0.0008700520265847445
step: 200, loss: 0.000210317550227046
step: 210, loss: 0.2134658396244049
step: 220, loss: 0.0007023180951364338
step: 230, loss: 0.009485973045229912
step: 240, loss: 0.0008545981254428625
step: 250, loss: 0.0003531741094775498
step: 260, loss: 0.0002708166721276939
step: 270, loss: 0.0005475524230860174
step: 280, loss: 0.0043219225481152534
step: 290, loss: 0.00021415407536551356
step: 300, loss: 0.0005435977363958955
step: 310, loss: 0.0002264254871988669
step: 320, loss: 0.0018783927662298083
step: 330, loss: 0.012347886338829994
step: 340, loss: 0.00011803188681369647
step: 350, loss: 9.811102063395083e-05
step: 360, loss: 0.00013735379616264254
step: 370, loss: 9.560152102494612e-05
step: 380, loss: 7.930156425572932e-05
step: 390, loss: 0.0033766361884772778
step: 400, loss: 0.00011506272858241573
step: 410, loss: 0.0011841236846521497
step: 420, loss: 0.0003844845632556826
step: 430, loss: 9.327563020633534e-05
step: 440, loss: 0.00023947504814714193
step: 450, loss: 0.0018828802276402712
step: 460, loss: 0.0003061036695726216
step: 470, loss: 0.19544294476509094
step: 480, loss: 0.0002683417114894837
step: 490, loss: 0.0018138671293854713
step: 500, loss: 0.000269322597887367
step: 510, loss: 0.01566426083445549
step: 520, loss: 0.0002592570090200752

```

step: 530, loss: 0.00044730110676027834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9388322520852641, f1=0.9249999999999999,
best_f1=0.9249999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006235329783521593
step: 10, loss: 0.0001290192740270868
step: 20, loss: 0.00020326831145212054
step: 30, loss: 0.0016332182567566633
step: 40, loss: 0.0008513341890648007
step: 50, loss: 0.0002372766175540164
step: 60, loss: 0.0005186967318877578
step: 70, loss: 0.00020238077559042722
step: 80, loss: 0.0004173356282990426
step: 90, loss: 0.0005922869895584881
step: 100, loss: 0.00039874581852927804
step: 110, loss: 0.0001144652851507999
step: 120, loss: 0.00019314035307615995
step: 130, loss: 0.0001900486822705716
step: 140, loss: 0.00023407266417052597
step: 150, loss: 0.0002932953939307481
step: 160, loss: 0.0007534656906500459
step: 170, loss: 0.0009350960026495159
step: 180, loss: 0.000229288125410676
step: 190, loss: 0.000176691755768843
step: 200, loss: 0.00017433628090657294
step: 210, loss: 0.000383699283702299
step: 220, loss: 0.00022798575810156763
step: 230, loss: 0.00036324639222584665
step: 240, loss: 0.0019160595256835222
step: 250, loss: 0.0002341000799788162
step: 260, loss: 0.005296738818287849
step: 270, loss: 0.00014560521231032908
step: 280, loss: 0.00012376521772239357
step: 290, loss: 0.0002380291116423905
step: 300, loss: 0.011856806464493275
step: 310, loss: 0.00021785467106383294
step: 320, loss: 0.0008445764542557299
step: 330, loss: 0.005427011288702488
step: 340, loss: 0.0008166056359186769

```

```

step: 350, loss: 9.491860691923648e-05
step: 360, loss: 0.18055446445941925
step: 370, loss: 0.00016052761930041015
step: 380, loss: 0.0003646781842689961
step: 390, loss: 0.0007854369468986988
step: 400, loss: 0.0011666918871924281
step: 410, loss: 0.0002533879305701703
step: 420, loss: 0.0020215415861457586
step: 430, loss: 0.0006112311384640634
step: 440, loss: 0.00010240777191938832
step: 450, loss: 0.00023338156461250037
step: 460, loss: 0.0007084052776917815
step: 470, loss: 0.0006368159083649516
step: 480, loss: 6.53291936032474e-05
step: 490, loss: 0.0003329381288494915
step: 500, loss: 7.679233385715634e-05
step: 510, loss: 0.00015288320719264448
step: 520, loss: 0.00022909812105353922
step: 530, loss: 0.0001130021337303333
/usr/local/lib/python3.7/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=0.9362101313320826, f1=0.9260299625468165,
best_f1=0.9249999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014158028352539986
step: 10, loss: 0.0001469468988943845
step: 20, loss: 0.0003451584780123085
step: 30, loss: 0.000182838921318762
step: 40, loss: 0.00022610688756685704
step: 50, loss: 0.00021711159206461161
step: 60, loss: 0.00020830084395129234
step: 70, loss: 0.016544196754693985
step: 80, loss: 0.0002395837800577283
step: 90, loss: 0.00010191415640292689
step: 100, loss: 0.00033634461578913033
step: 110, loss: 0.0003426242037676275
step: 120, loss: 0.00016032136045396328
step: 130, loss: 0.00013894654694013298
step: 140, loss: 0.0006722360267303884
step: 150, loss: 0.0002305235539097339
step: 160, loss: 0.0002525142044760287
step: 170, loss: 0.0016920791240409017
step: 180, loss: 0.00018870811618398875
step: 190, loss: 0.0014142110012471676
step: 200, loss: 0.0001960785302799195

```

```

step: 210, loss: 0.0014376454055309296
step: 220, loss: 9.686304110800847e-05
step: 230, loss: 0.0001399132888764143
step: 240, loss: 0.00027770467568188906
step: 250, loss: 0.0017045291606336832
step: 260, loss: 0.00014830603322479874
step: 270, loss: 0.00018181526684202254
step: 280, loss: 0.00010143908730242401
step: 290, loss: 6.0927348386030644e-05
step: 300, loss: 0.00036627482040785253
step: 310, loss: 5.810282891616225e-05
step: 320, loss: 0.006063106469810009
step: 330, loss: 0.0002881695982068777
step: 340, loss: 0.0002984104794450104
step: 350, loss: 9.992378909373656e-05
step: 360, loss: 0.0002356457116547972
step: 370, loss: 0.054162316024303436
step: 380, loss: 0.00039020032272674143
step: 390, loss: 0.0006341909756883979
step: 400, loss: 0.0011924817226827145
step: 410, loss: 7.343998004216701e-05
step: 420, loss: 0.00013187369040679187
step: 430, loss: 0.0005363732343539596
step: 440, loss: 0.0006225177785381675
step: 450, loss: 0.00032518056104891
step: 460, loss: 0.001063942676410079
step: 470, loss: 0.00044506846461445093
step: 480, loss: 0.00014326354721561074
step: 490, loss: 0.0002707403909880668
step: 500, loss: 0.21864920854568481
step: 510, loss: 0.00018940464360639453
step: 520, loss: 0.00014847413694951683
step: 530, loss: 0.000530216668266803
/usr/local/lib/python3.7/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=0.9382830626450116, f1=0.9302973977695168,
best_f1=0.9249999999999999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015375838847830892
step: 10, loss: 0.00028071634005755186
step: 20, loss: 0.0010049989214166999
step: 30, loss: 0.0022159989457577467
step: 40, loss: 0.00025003316113725305
step: 50, loss: 0.0002248992386739701
step: 60, loss: 0.00032764492789283395

```


step: 70, loss: 9.918045543599874e-05
step: 80, loss: 0.00014947287854738533
step: 90, loss: 0.000284438778180629
step: 100, loss: 0.0001560462696943432
step: 110, loss: 0.00023156352108344436
step: 120, loss: 0.00020860157383140177
step: 130, loss: 0.00011205138434888795
step: 140, loss: 0.00010673079668777063
step: 150, loss: 0.0001540264638606459
step: 160, loss: 0.00014107627794146538
step: 170, loss: 6.027829658705741e-05
step: 180, loss: 0.002787384670227766
step: 190, loss: 0.0010577145731076598
step: 200, loss: 0.00014256716531235725
step: 210, loss: 6.840316927991807e-05
step: 220, loss: 0.0001467294932808727
step: 230, loss: 0.22284908592700958
step: 240, loss: 0.0001376079162582755
step: 250, loss: 9.967236110242084e-05
step: 260, loss: 0.00011854703916469589
step: 270, loss: 0.000500204972922802
step: 280, loss: 0.00011672227992676198
step: 290, loss: 0.00025499192997813225
step: 300, loss: 0.00010245169687550515
step: 310, loss: 0.0001654848747421056
step: 320, loss: 0.0001717273407848552
step: 330, loss: 0.0012912803795188665
step: 340, loss: 0.0001723509922157973
step: 350, loss: 0.0004248374898452312
step: 360, loss: 0.009961182251572609
step: 370, loss: 0.00021143113553989679
step: 380, loss: 0.00016231724293902516
step: 390, loss: 0.0001577324146637693
step: 400, loss: 0.0011097162496298552
step: 410, loss: 0.0002619009464979172
step: 420, loss: 0.00013699823466595262
step: 430, loss: 8.010712917894125e-05
step: 440, loss: 0.022492432966828346
step: 450, loss: 0.0002730783016886562
step: 460, loss: 0.00013810826931148767
step: 470, loss: 0.0003547675732988864
step: 480, loss: 0.0006968821980990469
step: 490, loss: 0.0002753751177806407
step: 500, loss: 0.010031540878117085
step: 510, loss: 9.960844909073785e-05
step: 520, loss: 0.00014234479749575257
step: 530, loss: 0.00011953710054513067
/usr/local/lib/python3.7/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=0.9381682938168294, f1=0.9297347603536529,
best_f1=0.9249999999999999
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 [00:22, 255.14it/s]
load_f1 = 0.9362907031618688
real_f1 = 0.9327058823529412
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:20, 213.83it/s]

```

3.1.4 Amazon-Google - Running the matcher

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

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

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

```
Downloading: 100% 481/481 [00:00<00:00, 626kB/s]
Downloading: 100% 899k/899k [00:00<00:00, 41.1MB/s]
Downloading: 100% 456k/456k [00:00<00:00, 27.1MB/s]
Downloading: 100% 501M/501M [00:08<00:00, 61.8MB/s]
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.
```

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

```

master_weights      : True
loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4723604917526245
step: 10, loss: 0.4304710030555725
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.43054988980293274
step: 30, loss: 0.37666183710098267
step: 40, loss: 0.3559582233428955
step: 50, loss: 0.4335707724094391
step: 60, loss: 0.5193862318992615
step: 70, loss: 0.325621098279953
step: 80, loss: 0.3900032937526703
step: 90, loss: 0.3044467568397522
step: 100, loss: 0.26332470774650574
step: 110, loss: 0.2755528390407562
step: 120, loss: 0.4257674217224121
step: 130, loss: 0.22427153587341309
step: 140, loss: 0.45159465074539185
step: 150, loss: 0.394521564245224
step: 160, loss: 0.4781840443611145
step: 170, loss: 0.2379453331232071
step: 180, loss: 0.3999466896057129

```

```

step: 190, loss: 0.6294710636138916
step: 200, loss: 0.37338340282440186
step: 210, loss: 0.5120723247528076
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3682137131690979
step: 10, loss: 0.19759532809257507
step: 20, loss: 0.5242725610733032
step: 30, loss: 0.4936849772930145
step: 40, loss: 0.4614982306957245
step: 50, loss: 0.23269827663898468
step: 60, loss: 0.31792837381362915
step: 70, loss: 0.43318405747413635
step: 80, loss: 0.3166216015815735
step: 90, loss: 0.376044899225235
step: 100, loss: 0.525455892086029
step: 110, loss: 0.4003477692604065
step: 120, loss: 0.23708359897136688
step: 130, loss: 0.16913147270679474
step: 140, loss: 0.22191300988197327
step: 150, loss: 0.42380237579345703
step: 160, loss: 0.14513243734836578
step: 170, loss: 0.5383549332618713
step: 180, loss: 0.35108682513237
step: 190, loss: 0.31859248876571655
step: 200, loss: 0.13752807676792145
step: 210, loss: 0.3411328196525574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2850416302680969
step: 10, loss: 0.2430819422006607
step: 20, loss: 0.47222158312797546

```

```

step: 30, loss: 0.2720026671886444
step: 40, loss: 0.4515448808670044
step: 50, loss: 0.4626781940460205
step: 60, loss: 0.4895934462547302
step: 70, loss: 0.1745324283838272
step: 80, loss: 0.4909784495830536
step: 90, loss: 0.25392356514930725
step: 100, loss: 0.36883828043937683
step: 110, loss: 0.2313239872455597
step: 120, loss: 0.23339487612247467
step: 130, loss: 0.1578003615140915
step: 140, loss: 0.3788661062717438
step: 150, loss: 0.31403082609176636
step: 160, loss: 0.20573914051055908
step: 170, loss: 0.39430293440818787
step: 180, loss: 0.24889756739139557
step: 190, loss: 0.16210216283798218
step: 200, loss: 0.23030376434326172
step: 210, loss: 0.2614881694316864
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 3: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3308877646923065
step: 10, loss: 0.3052481412887573
step: 20, loss: 0.3109497129917145
step: 30, loss: 0.23608100414276123
step: 40, loss: 0.2411547303199768
step: 50, loss: 0.24495677649974823
step: 60, loss: 0.4811515510082245
step: 70, loss: 0.25202009081840515
step: 80, loss: 0.24603474140167236
step: 90, loss: 0.4434071183204651
step: 100, loss: 0.37829700112342834
step: 110, loss: 0.6003188490867615
step: 120, loss: 0.38540202379226685
step: 130, loss: 0.6806856393814087
step: 140, loss: 0.49365273118019104
step: 150, loss: 0.38724541664123535
step: 160, loss: 0.3150523900985718
step: 170, loss: 0.17197838425636292
step: 180, loss: 0.08194691687822342
step: 190, loss: 0.17282108962535858
step: 200, loss: 0.31506192684173584

```

```

step: 210, loss: 0.3871901333332062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3833116292953491
step: 10, loss: 0.3218291401863098
step: 20, loss: 0.31340131163597107
step: 30, loss: 0.24155253171920776
step: 40, loss: 0.4372464716434479
step: 50, loss: 0.3796613812446594
step: 60, loss: 0.38145601749420166
step: 70, loss: 0.2320481687784195
step: 80, loss: 0.4502662122249603
step: 90, loss: 0.44950759410858154
step: 100, loss: 0.24897059798240662
step: 110, loss: 0.16006232798099518
step: 120, loss: 0.23604242503643036
step: 130, loss: 0.3118494749069214
step: 140, loss: 0.5288630723953247
step: 150, loss: 0.31718695163726807
step: 160, loss: 0.24615810811519623
step: 170, loss: 0.3854867219924927
step: 180, loss: 0.23780867457389832
step: 190, loss: 0.5242255330085754
step: 200, loss: 0.3750189244747162
step: 210, loss: 0.24113129079341888
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18030251562595367
step: 10, loss: 0.37001630663871765
step: 20, loss: 0.31286612153053284
step: 30, loss: 0.24797922372817993
step: 40, loss: 0.30711403489112854
step: 50, loss: 0.6098356246948242
step: 60, loss: 0.3074911832809448
step: 70, loss: 0.43939635157585144
step: 80, loss: 0.2970495820045471

```

```

step: 90, loss: 0.3095347285270691
step: 100, loss: 0.3164246082305908
step: 110, loss: 0.39015117287635803
step: 120, loss: 0.25383666157722473
step: 130, loss: 0.24807430803775787
step: 140, loss: 0.39796701073646545
step: 150, loss: 0.19126775860786438
step: 160, loss: 0.16450154781341553
step: 170, loss: 0.45448943972587585
step: 180, loss: 0.37428995966911316
step: 190, loss: 0.44216763973236084
step: 200, loss: 0.31360405683517456
step: 210, loss: 0.3754833936691284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.45421814918518066
step: 10, loss: 0.2968152165412903
step: 20, loss: 0.43327197432518005
step: 30, loss: 0.25251248478889465
step: 40, loss: 0.24736814200878143
step: 50, loss: 0.3780701160430908
step: 60, loss: 0.3153366148471832
step: 70, loss: 0.2506773769855499
step: 80, loss: 0.5109924077987671
step: 90, loss: 0.3832811415195465
step: 100, loss: 0.4397260546684265
step: 110, loss: 0.3088163137435913
step: 120, loss: 0.31548917293548584
step: 130, loss: 0.471363365650177
step: 140, loss: 0.3108009099960327
step: 150, loss: 0.30874431133270264
step: 160, loss: 0.5365762710571289
step: 170, loss: 0.6346709728240967
step: 180, loss: 0.24692243337631226
step: 190, loss: 0.2377167046070099
step: 200, loss: 0.30960407853126526
step: 210, loss: 0.31973686814308167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6169945001602173
step: 10, loss: 0.3186132609844208
step: 20, loss: 0.2652955949306488
step: 30, loss: 0.24508114159107208
step: 40, loss: 0.24415740370750427
step: 50, loss: 0.17292919754981995
step: 60, loss: 0.17279286682605743
step: 70, loss: 0.4543134272098541
step: 80, loss: 0.31644493341445923
step: 90, loss: 0.36939379572868347
step: 100, loss: 0.6067623496055603
step: 110, loss: 0.3896312117576599
step: 120, loss: 0.30769357085227966
step: 130, loss: 0.16874469816684723
step: 140, loss: 0.31164032220840454
step: 150, loss: 0.4635770618915558
step: 160, loss: 0.45656558871269226
step: 170, loss: 0.524874210357666
step: 180, loss: 0.3099824786186218
step: 190, loss: 0.23211321234703064
step: 200, loss: 0.4592163562774658
step: 210, loss: 0.3807610273361206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5764499306678772
step: 10, loss: 0.2736869752407074
step: 20, loss: 0.37176644802093506
step: 30, loss: 0.16956345736980438
step: 40, loss: 0.16159170866012573
step: 50, loss: 0.4586215317249298
step: 60, loss: 0.17594748735427856
step: 70, loss: 0.3830184042453766
step: 80, loss: 0.32256221771240234
step: 90, loss: 0.31204473972320557
step: 100, loss: 0.4634765386581421
step: 110, loss: 0.5209469199180603
step: 120, loss: 0.3772510290145874
step: 130, loss: 0.23866526782512665
step: 140, loss: 0.5877878069877625

```

```

step: 150, loss: 0.121940016746521
step: 160, loss: 0.384701132774353
step: 170, loss: 0.3133842349052429
step: 180, loss: 0.4540957510471344
step: 190, loss: 0.31064844131469727
step: 200, loss: 0.3147696852684021
step: 210, loss: 0.31526485085487366
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3114335536956787
step: 10, loss: 0.3785233199596405
step: 20, loss: 0.2569655478000641
step: 30, loss: 0.16964951157569885
step: 40, loss: 0.3146226108074188
step: 50, loss: 0.5039660334587097
step: 60, loss: 0.2463492751121521
step: 70, loss: 0.38180238008499146
step: 80, loss: 0.16303236782550812
step: 90, loss: 0.38854414224624634
step: 100, loss: 0.4547804892063141
step: 110, loss: 0.17400118708610535
step: 120, loss: 0.45383691787719727
step: 130, loss: 0.2441311478614807
step: 140, loss: 0.38071611523628235
step: 150, loss: 0.2519019842147827
step: 160, loss: 0.3773283064365387
step: 170, loss: 0.24600493907928467
step: 180, loss: 0.3147714138031006
step: 190, loss: 0.2380709946155548
step: 200, loss: 0.6562443971633911
step: 210, loss: 0.25391069054603577
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.309752494096756
step: 10, loss: 0.31729063391685486
step: 20, loss: 0.379136860370636

```

```
step: 30, loss: 0.17055431008338928
step: 40, loss: 0.4560791254043579
step: 50, loss: 0.4378022253513336
step: 60, loss: 0.44887444376945496
step: 70, loss: 0.17252065241336823
step: 80, loss: 0.5276632905006409
step: 90, loss: 0.3743063807487488
step: 100, loss: 0.5124719738960266
step: 110, loss: 0.44321081042289734
step: 120, loss: 0.43257004022598267
step: 130, loss: 0.24898450076580048
step: 140, loss: 0.3219008147716522
step: 150, loss: 0.3110612630844116
step: 160, loss: 0.1681905835866928
step: 170, loss: 0.311848908662796
step: 180, loss: 0.2385735809803009
step: 190, loss: 0.44766825437545776
step: 200, loss: 0.18052417039871216
step: 210, loss: 0.3869333863258362
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24250636994838715
step: 10, loss: 0.32618263363838196
step: 20, loss: 0.4463657736778259
step: 30, loss: 0.32167285680770874
step: 40, loss: 0.177677184343338
step: 50, loss: 0.4642024636268616
step: 60, loss: 0.10839825868606567
step: 70, loss: 0.4526368975639343
step: 80, loss: 0.3164482116699219
step: 90, loss: 0.384159117937088
step: 100, loss: 0.1201367974281311
step: 110, loss: 0.25119978189468384
step: 120, loss: 0.2518548369407654
step: 130, loss: 0.43726998567581177
step: 140, loss: 0.3905302882194519
step: 150, loss: 0.10461939871311188
step: 160, loss: 0.24594029784202576
step: 170, loss: 0.2404123693704605
step: 180, loss: 0.310908704996109
step: 190, loss: 0.37438997626304626
step: 200, loss: 0.17679236829280853
```

```

step: 210, loss: 0.31195423007011414
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3105607330799103
step: 10, loss: 0.24109108746051788
step: 20, loss: 0.31492120027542114
step: 30, loss: 0.24560169875621796
step: 40, loss: 0.2486899346113205
step: 50, loss: 0.24327769875526428
step: 60, loss: 0.5772477388381958
step: 70, loss: 0.4424489736557007
step: 80, loss: 0.3772781491279602
step: 90, loss: 0.2406550496816635
step: 100, loss: 0.387044757604599
step: 110, loss: 0.24521638453006744
step: 120, loss: 0.30779051780700684
step: 130, loss: 0.3190784156322479
step: 140, loss: 0.1623355597257614
step: 150, loss: 0.3078306019306183
step: 160, loss: 0.5493983626365662
step: 170, loss: 0.23894605040550232
step: 180, loss: 0.24261705577373505
step: 190, loss: 0.2447420060634613
step: 200, loss: 0.3721892237663269
step: 210, loss: 0.3799295425415039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24763526022434235
step: 10, loss: 0.38501712679862976
step: 20, loss: 0.5030198097229004
step: 30, loss: 0.24767768383026123
step: 40, loss: 0.3103160858154297
step: 50, loss: 0.2494354546070099
step: 60, loss: 0.38298550248146057
step: 70, loss: 0.2488422393798828
step: 80, loss: 0.3126703202724457

```

```

step: 90, loss: 0.1656987965106964
step: 100, loss: 0.2467321902513504
step: 110, loss: 0.4519009590148926
step: 120, loss: 0.2453806847333908
step: 130, loss: 0.31417179107666016
step: 140, loss: 0.31918513774871826
step: 150, loss: 0.30793631076812744
step: 160, loss: 0.17462775111198425
step: 170, loss: 0.45499178767204285
step: 180, loss: 0.23918461799621582
step: 190, loss: 0.3711497485637665
step: 200, loss: 0.17144501209259033
step: 210, loss: 0.4470568299293518
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24084824323654175
step: 10, loss: 0.24403275549411774
step: 20, loss: 0.5208642482757568
step: 30, loss: 0.24484167993068695
step: 40, loss: 0.3764853775501251
step: 50, loss: 0.17043940722942352
step: 60, loss: 0.3141552209854126
step: 70, loss: 0.447885662317276
step: 80, loss: 0.2441485971212387
step: 90, loss: 0.4568396806716919
step: 100, loss: 0.31156033277511597
step: 110, loss: 0.30907177925109863
step: 120, loss: 0.44351252913475037
step: 130, loss: 0.17415347695350647
step: 140, loss: 0.3133874535560608
step: 150, loss: 0.519874095916748
step: 160, loss: 0.3800473213195801
step: 170, loss: 0.3804018795490265
step: 180, loss: 0.3153533339500427
step: 190, loss: 0.3086782991886139
step: 200, loss: 0.31611499190330505
step: 210, loss: 0.4524657428264618
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 477.78it/s]
load_f1 = 0.18519984170953702
real_f1 = 0.18519984170953702
267it [00:00, 1284.96it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 201.35it/s]
```

3.1.5 Walmart-Amazon - Running the matcher

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

```

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

```

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

```
!mv *_dev.pt checkpoints/Walmart-Amazon.pt
```

Step 3: run the matcher

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

```

Downloading: 100% 481/481 [00:00<00:00, 434kB/s]

Downloading: 100% 899k/899k [00:01<00:00, 745kB/s]

Downloading: 100% 456k/456k [00:00<00:00, 511kB/s]

Downloading: 100% 501M/501M [00:07<00:00, 69.9MB/s]

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4736657738685608
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4085082411766052
step: 20, loss: 0.35495859384536743
step: 30, loss: 0.36131083965301514
step: 40, loss: 0.23235072195529938
step: 50, loss: 0.3212697207927704
step: 60, loss: 0.44941744208335876
step: 70, loss: 0.4740677773952484
step: 80, loss: 0.174323171377182
step: 90, loss: 0.2898094356060028
step: 100, loss: 0.4342058300971985
step: 110, loss: 0.23106907308101654
step: 120, loss: 0.30908843874931335
step: 130, loss: 0.31253665685653687
step: 140, loss: 0.16675825417041779
step: 150, loss: 0.31390178203582764
step: 160, loss: 0.23104630410671234
step: 170, loss: 0.3856554627418518
step: 180, loss: 0.141786590218544
step: 190, loss: 0.16678209602832794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2692750287686997, f1=0.2883355176933159,

```



```

best_f1=0.2883355176933159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33381029963493347
step: 10, loss: 0.3284919559955597
step: 20, loss: 0.6651467680931091
step: 30, loss: 0.19048231840133667
step: 40, loss: 0.282368928194046
step: 50, loss: 0.35418376326560974
step: 60, loss: 0.355333536863327
step: 70, loss: 0.19042618572711945
step: 80, loss: 0.09391983598470688
step: 90, loss: 0.17782706022262573
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.18316057324409485
step: 110, loss: 0.14256636798381805
step: 120, loss: 0.19337785243988037
step: 130, loss: 0.1519353687763214
step: 140, loss: 0.21216394007205963
step: 150, loss: 0.2837497293949127
step: 160, loss: 0.24857476353645325
step: 170, loss: 0.03612251579761505
step: 180, loss: 0.09905631095170975
step: 190, loss: 0.16980868577957153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7486033519553073, f1=0.7878787878787877,
best_f1=0.7878787878787877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20341043174266815
step: 10, loss: 0.11474420130252838
step: 20, loss: 0.07725691050291061
step: 30, loss: 0.13504382967948914
step: 40, loss: 0.031163403764367104
step: 50, loss: 0.2507587969303131
step: 60, loss: 0.059484995901584625
step: 70, loss: 0.19424034655094147
step: 80, loss: 0.10717833042144775
step: 90, loss: 0.07280588150024414
step: 100, loss: 0.01793050579726696

```

```

step: 110, loss: 0.09935293346643448
step: 120, loss: 0.2743418514728546
step: 130, loss: 0.06168784201145172
step: 140, loss: 0.0624723918735981
step: 150, loss: 0.3269207179546356
step: 160, loss: 0.09982187300920486
step: 170, loss: 0.33456340432167053
step: 180, loss: 0.20292919874191284
step: 190, loss: 0.09219623357057571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7675070028011204, f1=0.8108108108108106,
best_f1=0.8108108108108106
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025134002789855003
step: 10, loss: 0.14403624832630157
step: 20, loss: 0.20042963325977325
step: 30, loss: 0.05865662172436714
step: 40, loss: 0.037064291536808014
step: 50, loss: 0.1227351576089859
step: 60, loss: 0.23245316743850708
step: 70, loss: 0.07458916306495667
step: 80, loss: 0.014026151970028877
step: 90, loss: 0.01619255729019642
step: 100, loss: 0.2082766890525818
step: 110, loss: 0.22759920358657837
step: 120, loss: 0.03896089270710945
step: 130, loss: 0.058867428451776505
step: 140, loss: 0.19797223806381226
step: 150, loss: 0.12348150461912155
step: 160, loss: 0.0392812043428421
step: 170, loss: 0.027662910521030426
step: 180, loss: 0.04321730509400368
step: 190, loss: 0.01466427929699421
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7988505747126435, f1=0.8076923076923078,
best_f1=0.8076923076923078
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11464536190032959
step: 10, loss: 0.1208314597606659
step: 20, loss: 0.06306721270084381
step: 30, loss: 0.07997939735651016
step: 40, loss: 0.06633511930704117
step: 50, loss: 0.049795884639024734
step: 60, loss: 0.014018640853464603
step: 70, loss: 0.2184801548719406
step: 80, loss: 0.013779952190816402
step: 90, loss: 0.015766389667987823
step: 100, loss: 0.21053196489810944
step: 110, loss: 0.30132782459259033
step: 120, loss: 0.013947729952633381
step: 130, loss: 0.20224568247795105
step: 140, loss: 0.12701238691806793
step: 150, loss: 0.11470449715852737
step: 160, loss: 0.01145044807344675
step: 170, loss: 0.05347111448645592
step: 180, loss: 0.05087735876441002
step: 190, loss: 0.018334561958909035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8328767123287673, f1=0.8074866310160429,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024010559543967247
step: 10, loss: 0.016228104010224342
step: 20, loss: 0.12622548639774323
step: 30, loss: 0.08438555896282196
step: 40, loss: 0.06283234059810638
step: 50, loss: 0.11462833732366562
step: 60, loss: 0.053619153797626495
step: 70, loss: 0.11331823468208313
step: 80, loss: 0.13735131919384003
step: 90, loss: 0.0053710429929196835
step: 100, loss: 0.15499262511730194

```

```

step: 110, loss: 0.012166688218712807
step: 120, loss: 0.030143234878778458
step: 130, loss: 0.158427894115448
step: 140, loss: 0.021713344380259514
step: 150, loss: 0.008595481514930725
step: 160, loss: 0.2286728173494339
step: 170, loss: 0.17419247329235077
step: 180, loss: 0.032903365790843964
step: 190, loss: 0.10352245718240738
/usr/local/lib/python3.7/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=0.8068181818181819, f1=0.8166666666666668,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008768647909164429
step: 10, loss: 0.18515458703041077
step: 20, loss: 0.006245137192308903
step: 30, loss: 0.010410208255052567
step: 40, loss: 0.00821542926132679
step: 50, loss: 0.002178728813305497
step: 60, loss: 0.009461804293096066
step: 70, loss: 0.003111285623162985
step: 80, loss: 0.0030664110090583563
step: 90, loss: 0.0045376247726380825
step: 100, loss: 0.18664434552192688
step: 110, loss: 0.16723008453845978
step: 120, loss: 0.006188520230352879
step: 130, loss: 0.04319660738110542
step: 140, loss: 0.028423329815268517
step: 150, loss: 0.1314331442117691
step: 160, loss: 0.15138383209705353
step: 170, loss: 0.021548660472035408
step: 180, loss: 0.0031084895599633455
step: 190, loss: 0.02814135141670704
/usr/local/lib/python3.7/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=0.8177083333333333, f1=0.8134715025906736,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024736231192946434
step: 10, loss: 0.0895603820681572
step: 20, loss: 0.017697656527161598

```

```

step: 30, loss: 0.18115650117397308
step: 40, loss: 0.03222188353538513
step: 50, loss: 0.018970727920532227
step: 60, loss: 0.01424458809196949
step: 70, loss: 0.009414112195372581
step: 80, loss: 0.18102623522281647
step: 90, loss: 0.02005523070693016
step: 100, loss: 0.0034087689127773046
step: 110, loss: 0.005517825484275818
step: 120, loss: 0.019650451838970184
step: 130, loss: 0.0017661405727267265
step: 140, loss: 0.027942148968577385
step: 150, loss: 0.021203337237238884
step: 160, loss: 0.10993639379739761
step: 170, loss: 0.001746136462315917
step: 180, loss: 0.004134473390877247
step: 190, loss: 0.040374912321567535
/usr/local/lib/python3.7/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=0.8131868131868132, f1=0.8337874659400545,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012300764210522175
step: 10, loss: 0.0008387180278077722
step: 20, loss: 0.00477201584726572
step: 30, loss: 0.01484409999102354
step: 40, loss: 0.1323455572128296
step: 50, loss: 0.01828695647418499
step: 60, loss: 0.009224959649145603
step: 70, loss: 0.002782511757686734
step: 80, loss: 0.10829399526119232
step: 90, loss: 0.02100498043000698
step: 100, loss: 0.06484396010637283
step: 110, loss: 0.060074079781770706
step: 120, loss: 0.0064178030006587505
step: 130, loss: 0.022979464381933212
step: 140, loss: 0.23332694172859192
step: 150, loss: 0.003446411108598113
step: 160, loss: 0.002172360895201564
step: 170, loss: 0.003462093183770776
step: 180, loss: 0.008497827686369419
step: 190, loss: 0.0012969061499461532
/usr/local/lib/python3.7/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=0.8133704735376045, f1=0.8164383561643836,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018325246637687087
step: 10, loss: 0.0034202630631625652
step: 20, loss: 0.02107042819261551
step: 30, loss: 0.14762945473194122
step: 40, loss: 0.0028178489301353693
step: 50, loss: 0.03950141742825508
step: 60, loss: 0.0020388762932270765
step: 70, loss: 0.0009895694674924016
step: 80, loss: 0.017106425017118454
step: 90, loss: 0.2326725423336029
step: 100, loss: 0.15563692152500153
step: 110, loss: 0.08932171016931534
step: 120, loss: 0.01014491356909275
step: 130, loss: 0.0043565756641328335
step: 140, loss: 0.006224576383829117
step: 150, loss: 0.0061805471777915955
step: 160, loss: 0.0022909652907401323
step: 170, loss: 0.003336644731462002
step: 180, loss: 0.00859172735363245
step: 190, loss: 0.0008602842572145164
/usr/local/lib/python3.7/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=0.7979274611398963, f1=0.7948051948051948,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012975919060409069
step: 10, loss: 0.0007983638206496835
step: 20, loss: 0.0010858859168365598
step: 30, loss: 0.02250901237130165
step: 40, loss: 0.001056235283613205
step: 50, loss: 0.05928169935941696
step: 60, loss: 0.0036648146342486143
step: 70, loss: 0.0009906168561428785
step: 80, loss: 0.014123180881142616
step: 90, loss: 0.0011876261560246348
step: 100, loss: 0.0017928596353158355
step: 110, loss: 0.15407666563987732
step: 120, loss: 0.0020771503914147615
step: 130, loss: 0.015037139877676964
step: 140, loss: 0.001445940462872386

```

```

step: 150, loss: 0.007451294921338558
step: 160, loss: 0.0032779790926724672
step: 170, loss: 0.005708937533199787
step: 180, loss: 0.0054599991999566555
step: 190, loss: 0.0008602036978118122
/usr/local/lib/python3.7/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=0.8146067415730337, f1=0.8212290502793296,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024795904755592346
step: 10, loss: 0.0005833492032252252
step: 20, loss: 0.0011196195846423507
step: 30, loss: 0.0010567338904365897
step: 40, loss: 0.0009131000260822475
step: 50, loss: 0.0008715499425306916
step: 60, loss: 0.003699617926031351
step: 70, loss: 0.0006802097777836025
step: 80, loss: 0.028670301660895348
step: 90, loss: 0.00064591143745929
step: 100, loss: 0.002248275326564908
step: 110, loss: 0.005148060619831085
step: 120, loss: 0.003508793655782938
step: 130, loss: 0.002304184017702937
step: 140, loss: 0.0008571039070375264
step: 150, loss: 0.0770438089966774
step: 160, loss: 0.014126328751444817
step: 170, loss: 0.005217605736106634
step: 180, loss: 0.0006835493841208518
step: 190, loss: 0.0007546594133600593
/usr/local/lib/python3.7/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=0.7977528089887641, f1=0.7955182072829131,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008760781493037939
step: 10, loss: 0.0009288575383834541
step: 20, loss: 0.004408939275890589
step: 30, loss: 0.01668134331703186
step: 40, loss: 0.0016516755567863584
step: 50, loss: 0.0005622666794806719
step: 60, loss: 0.007465103175491095

```

```

step: 70, loss: 0.004080415703356266
step: 80, loss: 0.0004899904597550631
step: 90, loss: 0.000350973685272038
step: 100, loss: 0.00046523427590727806
step: 110, loss: 0.0028886510990560055
step: 120, loss: 0.011761216446757317
step: 130, loss: 0.0007159888045862317
step: 140, loss: 0.01663973741233349
step: 150, loss: 0.0004418087482918054
step: 160, loss: 0.00044074218021705747
step: 170, loss: 0.0006433703820221126
step: 180, loss: 0.0006991180125623941
step: 190, loss: 0.07778561115264893
/usr/local/lib/python3.7/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=0.8112676056338027, f1=0.8032786885245903,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00046489739906974137
step: 10, loss: 0.0007556556956842542
step: 20, loss: 0.000621790299192071
step: 30, loss: 0.0004516477056313306
step: 40, loss: 0.002439639065414667
step: 50, loss: 0.0009170313132926822
step: 60, loss: 0.04845133796334267
step: 70, loss: 0.0027953756507486105
step: 80, loss: 0.017444007098674774
step: 90, loss: 0.002158456714823842
step: 100, loss: 0.0027431505732238293
step: 110, loss: 0.0008435146883130074
step: 120, loss: 0.0010373006807640195
step: 130, loss: 0.0005055837100371718
step: 140, loss: 0.0006722526741214097
step: 150, loss: 0.0006186846876516938
step: 160, loss: 0.000573475263081491
step: 170, loss: 0.0009411073406226933
step: 180, loss: 0.000713622197508812
step: 190, loss: 0.00046299066161736846
/usr/local/lib/python3.7/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=0.8228571428571428, f1=0.8033240997229917,
best_f1=0.8074866310160429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028515136800706387
step: 10, loss: 0.00045080724521540105
step: 20, loss: 0.0069000557996332645
step: 30, loss: 0.004221953917294741
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 40, loss: 0.0008920878753997386
step: 50, loss: 0.0006394676747731864
step: 60, loss: 0.0009905170882120728
step: 70, loss: 0.02379794232547283
step: 80, loss: 0.0005881928955204785
step: 90, loss: 0.01291744876652956
step: 100, loss: 0.003521207720041275
step: 110, loss: 0.0005540518322959542
step: 120, loss: 0.0004679172416217625
step: 130, loss: 0.0016671048942953348
step: 140, loss: 0.0005427206051535904
step: 150, loss: 0.0005881825345568359
step: 160, loss: 0.0004776332061737776
step: 170, loss: 0.2097119837999344
step: 180, loss: 0.0004478819901123643
step: 190, loss: 0.0014939281390979886
/usr/local/lib/python3.7/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=0.8228571428571428, f1=0.8033240997229917,
best_f1=0.8074866310160429
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 [00:14, 145.82it/s]
load_f1 = 0.8355795148247979
real_f1 = 0.8201058201058201
733it [00:00, 3662.50it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:33, 133.23it/s]
```

3.2 DK TEXTUAL

3.2.1 Abt-Buy - Running the matcher

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

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

# Step 3: run the matcher
!CUDA_VISIBLE_DEVICES=0 python matcher.py \
    --task Textual/Abt-Buy \
```

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

step: 0, loss: 0.468228280544281

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

step: 10, loss: 0.43666876077652

step: 20, loss: 0.3280518054962158

```

step: 30, loss: 0.5055815577507019
step: 40, loss: 0.5471956729888916
step: 50, loss: 0.2987542450428009
step: 60, loss: 0.5374029278755188
step: 70, loss: 0.29869112372398376
step: 80, loss: 0.2860306203365326
step: 90, loss: 0.21064060926437378
step: 100, loss: 0.14774104952812195
step: 110, loss: 0.4137565493583679
step: 120, loss: 0.30103763937950134
step: 130, loss: 0.3141029477119446
step: 140, loss: 0.38507455587387085
step: 150, loss: 0.30632394552230835
step: 160, loss: 0.3819105923175812
step: 170, loss: 0.30066928267478943
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.23413762287756926, f1=0.22265625, best_f1=0.22265625
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3181728720664978
step: 10, loss: 0.449340283870697
step: 20, loss: 0.34861499071121216
step: 30, loss: 0.30304381251335144
step: 40, loss: 0.07790058851242065
step: 50, loss: 0.4233095049858093
step: 60, loss: 0.17766880989074707
step: 70, loss: 0.48833590745925903
step: 80, loss: 0.2310100942850113
step: 90, loss: 0.2483326941728592
step: 100, loss: 0.5337349772453308
step: 110, loss: 0.29286929965019226
step: 120, loss: 0.24221400916576385
step: 130, loss: 0.5264973640441895
step: 140, loss: 0.5018743872642517
step: 150, loss: 0.43808168172836304
step: 160, loss: 0.3976886570453644
step: 170, loss: 0.36525145173072815
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.24796380090497736, f1=0.2380038387715931,

best_f1=0.2380038387715931

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

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

step: 0, loss: 0.5825449228286743

step: 10, loss: 0.30311691761016846

step: 20, loss: 0.23282502591609955

step: 30, loss: 0.23110316693782806

step: 40, loss: 0.3562746047973633

step: 50, loss: 0.5810937285423279

step: 60, loss: 0.2751709222793579

step: 70, loss: 0.23460876941680908

step: 80, loss: 0.3921346068382263

step: 90, loss: 0.5278863906860352

step: 100, loss: 0.23720799386501312

step: 110, loss: 0.17311899363994598

step: 120, loss: 0.4793703854084015

step: 130, loss: 0.5074244737625122

step: 140, loss: 0.41755783557891846

step: 150, loss: 0.20421993732452393

step: 160, loss: 0.16590774059295654

step: 170, loss: 0.31160441040992737

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

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

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

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=0.3016393442622951, f1=0.2777777777777778,

best_f1=0.2777777777777778

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

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

step: 0, loss: 0.4115790128707886

step: 10, loss: 0.5140698552131653

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

step: 20, loss: 0.18008467555046082

step: 30, loss: 0.33277642726898193

step: 40, loss: 0.20545713603496552

step: 50, loss: 0.3300466537475586

step: 60, loss: 0.469350129365921

step: 70, loss: 0.2967468500137329

step: 80, loss: 0.4584999084472656

```

step: 90, loss: 0.3226306736469269
step: 100, loss: 0.26472437381744385
step: 110, loss: 0.4588603973388672
step: 120, loss: 0.47185686230659485
step: 130, loss: 0.15858419239521027
step: 140, loss: 0.5123539566993713
step: 150, loss: 0.9495685696601868
step: 160, loss: 0.17930968105793
step: 170, loss: 0.2668409049510956
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6099585062240663, f1=0.625531914893617,
best_f1=0.625531914893617
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3023938536643982
step: 10, loss: 0.21888715028762817
step: 20, loss: 0.18639828264713287
step: 30, loss: 0.07399938255548477
step: 40, loss: 0.18563951551914215
step: 50, loss: 0.11207620054483414
step: 60, loss: 0.30426672101020813
step: 70, loss: 0.09399683773517609
step: 80, loss: 0.03900960087776184
step: 90, loss: 0.2806570529937744
step: 100, loss: 0.07282810658216476
step: 110, loss: 0.2031085342168808
step: 120, loss: 0.09170851111412048
step: 130, loss: 0.03892055153846741
step: 140, loss: 0.09377388656139374
step: 150, loss: 0.12392992526292801
step: 160, loss: 0.2219681441783905
step: 170, loss: 0.14261257648468018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7707317073170732, f1=0.7554479418886199,
best_f1=0.7554479418886199

```

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

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

step: 0, loss: 0.1469840705394745
step: 10, loss: 0.03689025342464447
step: 20, loss: 0.2572741210460663
step: 30, loss: 0.01957692578434944
step: 40, loss: 0.07056240737438202
step: 50, loss: 0.18189674615859985
step: 60, loss: 0.11095859110355377
step: 70, loss: 0.08419103175401688
step: 80, loss: 0.08133908361196518
step: 90, loss: 0.18044336140155792
step: 100, loss: 0.05561079829931259
step: 110, loss: 0.20997938513755798
step: 120, loss: 0.2212788611650467
step: 130, loss: 0.1797635853290558
step: 140, loss: 0.18735532462596893
step: 150, loss: 0.25160637497901917
step: 160, loss: 0.2777596414089203
step: 170, loss: 0.060199007391929626

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

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

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

warnings.warn(SAVE_STATE_WARNING, UserWarning)

epoch 6: dev_f1=0.7855421686746988, f1=0.7625899280575539,
best_f1=0.7625899280575539

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

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

step: 0, loss: 0.13464604318141937
step: 10, loss: 0.13233570754528046
step: 20, loss: 0.2006557434797287
step: 30, loss: 0.29579171538352966
step: 40, loss: 0.022347569465637207
step: 50, loss: 0.0390123575925827
step: 60, loss: 0.10842785239219666
step: 70, loss: 0.03926295414566994
step: 80, loss: 0.10133066773414612
step: 90, loss: 0.05372961610555649
step: 100, loss: 0.01250184141099453
step: 110, loss: 0.045640040189027786
step: 120, loss: 0.039422258734703064
step: 130, loss: 0.10588866472244263
step: 140, loss: 0.04610278457403183

```

step: 150, loss: 0.13899827003479004
step: 160, loss: 0.03463472053408623
step: 170, loss: 0.05185304954648018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.804177545691906, f1=0.7589743589743588,
best_f1=0.7589743589743588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16688235104084015
step: 10, loss: 0.01678118109703064
step: 20, loss: 0.05986028164625168
step: 30, loss: 0.005538852885365486
step: 40, loss: 0.16515471041202545
step: 50, loss: 0.03052217699587345
step: 60, loss: 0.039443835616111755
step: 70, loss: 0.08631112426519394
step: 80, loss: 0.16179661452770233
step: 90, loss: 0.08749232441186905
step: 100, loss: 0.0066065313294529915
step: 110, loss: 0.10386821627616882
step: 120, loss: 0.36398571729660034
step: 130, loss: 0.027225086465477943
step: 140, loss: 0.0951484739780426
step: 150, loss: 0.022283442318439484
step: 160, loss: 0.008334673941135406
step: 170, loss: 0.15650033950805664
/usr/local/lib/python3.7/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=0.8030690537084398, f1=0.7660668380462724,
best_f1=0.7589743589743588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025973239913582802
step: 10, loss: 0.031425535678863525
step: 20, loss: 0.019521918147802353
step: 30, loss: 0.1347009390592575
step: 40, loss: 0.04238610342144966
step: 50, loss: 0.005855815950781107
step: 60, loss: 0.08086606115102768

```



```

step: 70, loss: 0.12749414145946503
step: 80, loss: 0.0034386315383017063
step: 90, loss: 0.08065260946750641
step: 100, loss: 0.09389431774616241
step: 110, loss: 0.037636347115039825
step: 120, loss: 0.018912039697170258
step: 130, loss: 0.23109892010688782
step: 140, loss: 0.03599195182323456
step: 150, loss: 0.2960360646247864
step: 160, loss: 0.03344593197107315
step: 170, loss: 0.03535836562514305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8073878627968338, f1=0.763157894736842,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13619881868362427
step: 10, loss: 0.051475949585437775
step: 20, loss: 0.1363130360841751
step: 30, loss: 0.08209428191184998
step: 40, loss: 0.026426002383232117
step: 50, loss: 0.060883667320013046
step: 60, loss: 0.12114103883504868
step: 70, loss: 0.011295230127871037
step: 80, loss: 0.028022238984704018
step: 90, loss: 0.004410641733556986
step: 100, loss: 0.022693810984492302
step: 110, loss: 0.012098686769604683
step: 120, loss: 0.018980061635375023
step: 130, loss: 0.022488849237561226
step: 140, loss: 0.006517288740724325
step: 150, loss: 0.023475343361496925
step: 160, loss: 0.11436014622449875
step: 170, loss: 0.060858502984046936
/usr/local/lib/python3.7/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=0.7748691099476439, f1=0.7611548556430446,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045883432030677795
step: 10, loss: 0.001960438210517168
step: 20, loss: 0.05860396474599838
step: 30, loss: 0.05608599632978439
step: 40, loss: 0.044749025255441666
step: 50, loss: 0.09258533269166946
step: 60, loss: 0.043990932404994965
step: 70, loss: 0.0014077863888815045
step: 80, loss: 0.007769694086164236
step: 90, loss: 0.08597040921449661
step: 100, loss: 0.25261324644088745
step: 110, loss: 0.0025732964277267456
step: 120, loss: 0.007524797692894936
step: 130, loss: 0.05444395914673805
step: 140, loss: 0.12824822962284088
step: 150, loss: 0.045370399951934814
step: 160, loss: 0.060900021344423294
step: 170, loss: 0.22085873782634735
/usr/local/lib/python3.7/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=0.7936507936507937, f1=0.7586206896551725,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005154763348400593
step: 10, loss: 0.03394082561135292
step: 20, loss: 0.01265487726777792
step: 30, loss: 0.06203009560704231
step: 40, loss: 0.040411315858364105
step: 50, loss: 0.0024050516076385975
step: 60, loss: 0.022480633109807968
step: 70, loss: 0.06194687262177467
step: 80, loss: 0.004075606353580952
step: 90, loss: 0.03447999060153961
step: 100, loss: 0.004478045739233494
step: 110, loss: 0.04883892461657524
step: 120, loss: 0.04550553485751152
step: 130, loss: 0.15509547293186188
step: 140, loss: 0.0031949239782989025
step: 150, loss: 0.0032273794058710337
step: 160, loss: 0.06558464467525482
step: 170, loss: 0.025851868093013763
/usr/local/lib/python3.7/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=0.7892156862745098, f1=0.7642679900744417,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005965011194348335
step: 10, loss: 0.024469301104545593
step: 20, loss: 0.006073683500289917
step: 30, loss: 0.030672328546643257
step: 40, loss: 0.02526368945837021
step: 50, loss: 0.035385895520448685
step: 60, loss: 0.004433070309460163
step: 70, loss: 0.03074328601360321
step: 80, loss: 0.0014039411908015609
step: 90, loss: 0.004171835258603096
step: 100, loss: 0.04052966833114624
step: 110, loss: 0.0018707021372392774
step: 120, loss: 0.005307129118591547
step: 130, loss: 0.0043645212426781654
step: 140, loss: 0.014512826688587666
step: 150, loss: 0.0014787460677325726
step: 160, loss: 0.028873061761260033
step: 170, loss: 0.0022331979125738144
/usr/local/lib/python3.7/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=0.7949367088607596, f1=0.7518796992481204,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016001543262973428
step: 10, loss: 0.004544997587800026
step: 20, loss: 0.0339859314262867
step: 30, loss: 0.03482525050640106
step: 40, loss: 0.0026799372863024473
step: 50, loss: 0.0019592507742345333
step: 60, loss: 0.009247456677258015
step: 70, loss: 0.00890295673161745
step: 80, loss: 0.006823115982115269
step: 90, loss: 0.005871148779988289
step: 100, loss: 0.0011255725985392928
step: 110, loss: 0.01429822389036417
step: 120, loss: 0.02722003310918808
step: 130, loss: 0.12842732667922974
step: 140, loss: 0.035603102296590805
step: 150, loss: 0.016346139833331108
step: 160, loss: 0.030978567898273468

```

```

step: 170, loss: 0.049193937331438065
/usr/local/lib/python3.7/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=0.7926509186351706, f1=0.7676240208877285,
best_f1=0.763157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004584414418786764
step: 10, loss: 0.01388260256499052
step: 20, loss: 0.22611476480960846
step: 30, loss: 0.0038309672381728888
step: 40, loss: 0.000684835365973413
step: 50, loss: 0.0007920420030131936
step: 60, loss: 0.0014965386362746358
step: 70, loss: 0.012725657783448696
step: 80, loss: 0.0016770090442150831
step: 90, loss: 0.0014160617720335722
step: 100, loss: 0.01596944034099579
step: 110, loss: 0.00986375194042921
step: 120, loss: 0.01196654699742794
step: 130, loss: 0.00647604838013649
step: 140, loss: 0.005365564487874508
step: 150, loss: 0.0045166234485805035
step: 160, loss: 0.030161703005433083
step: 170, loss: 0.0009560775943100452
/usr/local/lib/python3.7/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=0.7894736842105263, f1=0.763157894736842,
best_f1=0.763157894736842
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled          : True
opt_level        : 02

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:08, 219.95it/s]
load_f1 = 0.735576923076923
real_f1 = 0.7342995169082126
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.80it/s]

```

3.3 DK DIRTY - oK

3.3.1 DBLP-ACM - Running the matcher

```

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

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

```

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.6279656887054443
```

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

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

```
step: 10, loss: 0.42634278535842896
```

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

```
step: 20, loss: 0.4729517996311188
```

```
step: 30, loss: 0.35656365752220154
```

```
step: 40, loss: 0.3370693325996399
```

```
step: 50, loss: 0.7324949502944946
```

```

step: 60, loss: 0.46691861748695374
step: 70, loss: 0.41002601385116577
step: 80, loss: 0.5868573188781738
step: 90, loss: 0.2465859055519104
step: 100, loss: 0.5803751349449158
step: 110, loss: 0.5226289629936218
step: 120, loss: 0.4152146875858307
step: 130, loss: 0.2998940944671631
step: 140, loss: 0.3430609703063965
step: 150, loss: 0.5298281311988831
step: 160, loss: 0.20419323444366455
step: 170, loss: 0.3549209237098694
step: 180, loss: 0.26754534244537354
step: 190, loss: 0.08197126537561417
step: 200, loss: 0.20036497712135315
step: 210, loss: 0.3581770062446594
step: 220, loss: 0.16341249644756317
step: 230, loss: 0.07199973613023758
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8507936507936508, f1=0.8443496801705757,
best_f1=0.8443496801705757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06546354293823242
step: 10, loss: 0.31064102053642273
step: 20, loss: 0.1109025776386261
step: 30, loss: 0.2948839068412781
step: 40, loss: 0.13031907379627228
step: 50, loss: 0.050674185156822205
step: 60, loss: 0.02626366913318634
step: 70, loss: 0.10988323390483856
step: 80, loss: 0.040045663714408875
step: 90, loss: 0.08037838339805603
step: 100, loss: 0.050494227558374405
step: 110, loss: 0.16911578178405762
step: 120, loss: 0.04862217977643013
step: 130, loss: 0.07481884211301804
step: 140, loss: 0.04869983717799187
step: 150, loss: 0.17100916802883148
step: 160, loss: 0.13222377002239227
step: 170, loss: 0.004277760628610849

```

```

step: 180, loss: 0.033810943365097046
step: 190, loss: 0.019585955888032913
step: 200, loss: 0.15524275600910187
step: 210, loss: 0.0826081857085228
step: 220, loss: 0.1798919290304184
step: 230, loss: 0.004366443492472172
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9463307776560789, f1=0.9477196885428254,
best_f1=0.9477196885428254
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0854584127664566
step: 10, loss: 0.05610809102654457
step: 20, loss: 0.06441565603017807
step: 30, loss: 0.06069318950176239
step: 40, loss: 0.0822448581457138
step: 50, loss: 0.04026154801249504
step: 60, loss: 0.06267418712377548
step: 70, loss: 0.10610753297805786
step: 80, loss: 0.026298336684703827
step: 90, loss: 0.05787596106529236
step: 100, loss: 0.06695034354925156
step: 110, loss: 0.018618563190102577
step: 120, loss: 0.013139354065060616
step: 130, loss: 0.033633917570114136
step: 140, loss: 0.006797193083912134
step: 150, loss: 0.09905465692281723
step: 160, loss: 0.010056020691990852
step: 170, loss: 0.002867491915822029
step: 180, loss: 0.11032086610794067
step: 190, loss: 0.07696081697940826
step: 200, loss: 0.020102597773075104
step: 210, loss: 0.005994857754558325
step: 220, loss: 0.1224244013428688
step: 230, loss: 0.06527833640575409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9539473684210527, f1=0.9565217391304347,
best_f1=0.9565217391304347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09570752829313278
step: 10, loss: 0.01898021064698696
step: 20, loss: 0.0048052589409053326
step: 30, loss: 0.0019844730850309134
step: 40, loss: 0.13979817926883698
step: 50, loss: 0.026156999170780182
step: 60, loss: 0.05660145729780197
step: 70, loss: 0.03750486299395561
step: 80, loss: 0.015160778537392616
step: 90, loss: 0.09736104309558868
step: 100, loss: 0.0603974312543869
step: 110, loss: 0.007684402167797089
step: 120, loss: 0.02594120241701603
step: 130, loss: 0.006372493226081133
step: 140, loss: 0.005922677926719189
step: 150, loss: 0.010061549954116344
step: 160, loss: 0.012395060621201992
step: 170, loss: 0.0047273882664740086
step: 180, loss: 0.07563607394695282
step: 190, loss: 0.010304532945156097
step: 200, loss: 0.12211423367261887
step: 210, loss: 0.0025683243293315172
step: 220, loss: 0.0026776615995913744
step: 230, loss: 0.01056975033134222
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9665178571428571, f1=0.9580022701475595,
best_f1=0.9580022701475595
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002391661750152707
step: 10, loss: 0.036853838711977005
step: 20, loss: 0.07822806388139725
step: 30, loss: 0.013161142356693745
step: 40, loss: 0.026813147589564323
step: 50, loss: 0.03259899839758873

```

```

step: 60, loss: 0.09601069986820221
step: 70, loss: 0.01437241118401289
step: 80, loss: 0.05437208339571953
step: 90, loss: 0.05578872188925743
step: 100, loss: 0.0005210903473198414
step: 110, loss: 0.007141198497265577
step: 120, loss: 0.004636053461581469
step: 130, loss: 0.0022024260833859444
step: 140, loss: 0.1005476713180542
step: 150, loss: 0.021129127591848373
step: 160, loss: 0.0009738240041770041
step: 170, loss: 0.03885950148105621
step: 180, loss: 0.026109740138053894
step: 190, loss: 0.10844758152961731
step: 200, loss: 0.02646993286907673
step: 210, loss: 0.0064199939370155334
step: 220, loss: 0.02947630174458027
step: 230, loss: 0.015139181166887283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9723145071982282, f1=0.9708520179372198,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009292985196225345
step: 10, loss: 0.003285585204139352
step: 20, loss: 0.0032260995358228683
step: 30, loss: 0.0015355634968727827
step: 40, loss: 0.0005443824920803308
step: 50, loss: 0.014059459790587425
step: 60, loss: 0.12599456310272217
step: 70, loss: 0.11687352508306503
step: 80, loss: 0.034783292561769485
step: 90, loss: 0.0020995105151087046
step: 100, loss: 0.0015036994591355324
step: 110, loss: 0.009686626493930817
step: 120, loss: 0.0010818653972819448
step: 130, loss: 0.01531602069735527
step: 140, loss: 0.002923050429672003
step: 150, loss: 0.004311999771744013
step: 160, loss: 0.002631785813719034
step: 170, loss: 0.0014617444248870015

```

```

step: 180, loss: 0.003346384735777974
step: 190, loss: 0.005952255334705114
step: 200, loss: 0.002659567166119814
step: 210, loss: 0.002549707191064954
step: 220, loss: 0.020190274342894554
step: 230, loss: 0.0012219061609357595
/usr/local/lib/python3.7/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=0.9696287964004499, f1=0.963718820861678,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012066283961758018
step: 10, loss: 0.0005945658776909113
step: 20, loss: 0.003343098098412156
step: 30, loss: 0.0010912674479186535
step: 40, loss: 0.0014656024286523461
step: 50, loss: 0.004781703464686871
step: 60, loss: 0.003123482456430793
step: 70, loss: 0.0028012169059365988
step: 80, loss: 0.0012282554525882006
step: 90, loss: 0.0027250314597040415
step: 100, loss: 0.0005713102291338146
step: 110, loss: 0.0010042282519862056
step: 120, loss: 0.047766879200935364
step: 130, loss: 0.006839178502559662
step: 140, loss: 0.0010341006563976407
step: 150, loss: 0.2214994579553604
step: 160, loss: 0.0006542472983710468
step: 170, loss: 0.0015479158610105515
step: 180, loss: 0.0009091483661904931
step: 190, loss: 0.1117030680179596
step: 200, loss: 0.0021688782144337893
step: 210, loss: 0.006035075057297945
step: 220, loss: 0.002301272237673402
step: 230, loss: 0.0181965883821249
/usr/local/lib/python3.7/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=0.9367681498829039, f1=0.9223416965352449,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0422699861228466
step: 10, loss: 0.0038297686260193586

```

```

step: 20, loss: 0.006682452745735645
step: 30, loss: 0.001916907262057066
step: 40, loss: 0.0016315871616825461
step: 50, loss: 0.007233213633298874
step: 60, loss: 0.002180559327825904
step: 70, loss: 0.00045421053073368967
step: 80, loss: 0.0973396748304367
step: 90, loss: 0.005230179987847805
step: 100, loss: 0.011991791427135468
step: 110, loss: 0.009498367086052895
step: 120, loss: 0.01600363291800022
step: 130, loss: 0.005213124677538872
step: 140, loss: 0.00038863776717334986
step: 150, loss: 0.18016935884952545
step: 160, loss: 0.0008110290509648621
step: 170, loss: 0.009452472440898418
step: 180, loss: 0.0018649811390787363
step: 190, loss: 0.0056971898302435875
step: 200, loss: 0.024647993966937065
step: 210, loss: 0.003925316967070103
step: 220, loss: 0.0012995079159736633
step: 230, loss: 0.0008594793034717441
/usr/local/lib/python3.7/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=0.9698324022346367, f1=0.9578107183580388,
best_f1=0.9708520179372198
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013866517692804337
step: 10, loss: 0.0014505821745842695
step: 20, loss: 0.001164392801001668
step: 30, loss: 0.0038113927002996206
step: 40, loss: 0.16323935985565186
step: 50, loss: 0.001973165897652507
step: 60, loss: 0.0017517171800136566
step: 70, loss: 0.015086548402905464
step: 80, loss: 0.0033073099330067635
step: 90, loss: 0.14733503758907318
step: 100, loss: 0.0010427674278616905
step: 110, loss: 0.00035801157355308533
step: 120, loss: 0.025670507922768593
step: 130, loss: 0.00460497010499239
step: 140, loss: 0.012700100429356098
step: 150, loss: 0.0006334125646390021
step: 160, loss: 0.0717429667711258
step: 170, loss: 0.0004662818100769073

```

```

step: 180, loss: 0.0005455560167320073
step: 190, loss: 0.0006503271288238466
step: 200, loss: 0.0008850594749674201
step: 210, loss: 0.0007823788328096271
step: 220, loss: 0.0005180962034501135
step: 230, loss: 0.0005762330838479102
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9785794813979707, f1=0.9657534246575342,
best_f1=0.9657534246575342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008422807441093028
step: 10, loss: 0.0007891812128946185
step: 20, loss: 0.0004651197523344308
step: 30, loss: 0.00017096954979933798
step: 40, loss: 0.0011755269952118397
step: 50, loss: 0.00015926976629998535
step: 60, loss: 0.0002971486246678978
step: 70, loss: 0.00435241125524044
step: 80, loss: 0.00022342696320265532
step: 90, loss: 0.0004863193607889116
step: 100, loss: 0.00019206557772122324
step: 110, loss: 0.008193562738597393
step: 120, loss: 0.00010351950186304748
step: 130, loss: 0.01946563832461834
step: 140, loss: 0.0001299217256018892
step: 150, loss: 0.0019453297136351466
step: 160, loss: 8.542246359866112e-05
step: 170, loss: 0.0011699238093569875
step: 180, loss: 0.30117976665496826
step: 190, loss: 0.001144132693298161
step: 200, loss: 0.0005861440440639853
step: 210, loss: 0.0024800929240882397
step: 220, loss: 0.0005896310904063284
step: 230, loss: 0.00045891935587860644
/usr/local/lib/python3.7/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=0.9584295612009238,
best_f1=0.9657534246575342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027807449805550277
step: 10, loss: 0.0005928651662543416
step: 20, loss: 0.0002808299905154854
step: 30, loss: 0.0067019169218838215
step: 40, loss: 0.00023386995599139482
step: 50, loss: 0.0012310580350458622
step: 60, loss: 0.006241902709007263
step: 70, loss: 0.00044514486216939986
step: 80, loss: 0.0005327895632945001
step: 90, loss: 0.051564771682024
step: 100, loss: 0.0002673457784112543
step: 110, loss: 0.0011095245135948062
step: 120, loss: 0.00030277561745606363
step: 130, loss: 0.00012232545122969896
step: 140, loss: 0.002148005645722151
step: 150, loss: 0.0004951465525664389
step: 160, loss: 0.01055504847317934
step: 170, loss: 0.15333403646945953
step: 180, loss: 0.011918717063963413
step: 190, loss: 0.0006030109361745417
step: 200, loss: 0.008155248127877712
step: 210, loss: 0.0003229143039789051
step: 220, loss: 0.0009350280161015689
step: 230, loss: 0.0016741572180762887
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9788182831661093, f1=0.9662921348314607,
best_f1=0.9662921348314607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019302857108414173
step: 10, loss: 0.00012204437371110544
step: 20, loss: 0.0008045663707889616
step: 30, loss: 0.03476254269480705
step: 40, loss: 0.0002564018068369478
step: 50, loss: 0.0002776375040411949
step: 60, loss: 0.001690295641310513
step: 70, loss: 0.0001100774752558209
step: 80, loss: 8.514522050973028e-05
step: 90, loss: 0.006072851829230785

```

```

step: 100, loss: 0.00011668707884382457
step: 110, loss: 0.00014254178677219898
step: 120, loss: 0.00014982452557887882
step: 130, loss: 0.0002724990772549063
step: 140, loss: 0.00015461008297279477
step: 150, loss: 0.00020171959477011114
step: 160, loss: 0.0032661932054907084
step: 170, loss: 0.00029937236104160547
step: 180, loss: 0.00024001255223993212
step: 190, loss: 0.001276442315429449
step: 200, loss: 0.0005142696900293231
step: 210, loss: 0.0416165292263031
step: 220, loss: 9.975220746127889e-05
step: 230, loss: 0.0003066598728764802
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9799107142857142, f1=0.9628796400449944,
best_f1=0.9628796400449944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002902248816099018
step: 10, loss: 0.005607893690466881
step: 20, loss: 0.00016865527140907943
step: 30, loss: 6.440935248974711e-05
step: 40, loss: 0.0007117597269825637
step: 50, loss: 0.003916042856872082
step: 60, loss: 0.00034436071291565895
step: 70, loss: 0.0001663608563831076
step: 80, loss: 0.0037494597490876913
step: 90, loss: 0.000706656719557941
step: 100, loss: 0.0002615690464153886
step: 110, loss: 0.00046947653754614294
step: 120, loss: 0.00016317496192641556
step: 130, loss: 0.0006149573600850999
step: 140, loss: 0.00011790992721216753
step: 150, loss: 0.0002552381483837962
step: 160, loss: 0.01790563575923443
step: 170, loss: 0.00011068764433730394
step: 180, loss: 0.005467269569635391
step: 190, loss: 0.00010453631693962961
step: 200, loss: 3.469549847068265e-05
step: 210, loss: 0.00013047784159425646

```

```

step: 220, loss: 0.00028423607000149786
step: 230, loss: 0.0022362901363521814
/usr/local/lib/python3.7/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=0.9740112994350283, f1=0.9589041095890412,
best_f1=0.9628796400449944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000271655066171661
step: 10, loss: 0.00018895151151809841
step: 20, loss: 0.0005309656262397766
step: 30, loss: 0.0006464957259595394
step: 40, loss: 0.00021810246107634157
step: 50, loss: 8.758968760957941e-05
step: 60, loss: 0.00014510143955703825
step: 70, loss: 0.0007497541955672204
step: 80, loss: 0.0006265437114052474
step: 90, loss: 0.00028098930488340557
step: 100, loss: 0.00014191213995218277
step: 110, loss: 0.00021100160665810108
step: 120, loss: 5.384681571740657e-05
step: 130, loss: 0.0004032707947771996
step: 140, loss: 0.0003265418636146933
step: 150, loss: 0.0024371568579226732
step: 160, loss: 0.00024238074547611177
step: 170, loss: 0.00013297572149895132
step: 180, loss: 0.0009321335237473249
step: 190, loss: 8.737562166061252e-05
step: 200, loss: 0.0004519208159763366
step: 210, loss: 8.525617886334658e-05
step: 220, loss: 0.0002576721308287233
step: 230, loss: 6.50084184599109e-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=0.978675645342312, f1=0.9648924122310306,
best_f1=0.9628796400449944
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001308258855715394
step: 10, loss: 0.0002496731176506728
step: 20, loss: 0.0006910550873726606
step: 30, loss: 0.0001065947362803854
step: 40, loss: 6.775976362405345e-05
step: 50, loss: 0.00010572135943220928

```



```

step: 60, loss: 0.002353296149522066
step: 70, loss: 0.007429852616041899
step: 80, loss: 0.00016112883167807013
step: 90, loss: 8.649779192637652e-05
step: 100, loss: 0.0009044783073477447
step: 110, loss: 0.00011028223525499925
step: 120, loss: 0.03683643415570259
step: 130, loss: 0.0009000226855278015
step: 140, loss: 0.0017834256868809462
step: 150, loss: 0.00016608666919637471
step: 160, loss: 0.0006971480324864388
step: 170, loss: 8.785104000708088e-05
step: 180, loss: 0.00016429588140454143
step: 190, loss: 0.00035230661160312593
step: 200, loss: 0.0021040558349341154
step: 210, loss: 0.0005746794049628079
step: 220, loss: 7.825395732652396e-05
step: 230, loss: 9.827130270423368e-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=0.9764837625979844, f1=0.9638009049773756,
best_f1=0.9628796400449944
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
2473it [00:17, 142.20it/s]
load_f1 = 0.9821029082774049
real_f1 = 0.9788182831661093
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:32, 136.26it/s]
```

3.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

step: 0, loss: 0.6389544606208801

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

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

step: 10, loss: 0.5037962198257446

step: 20, loss: 0.32592684030532837

step: 30, loss: 0.42839083075523376

step: 40, loss: 0.32375508546829224

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

step: 50, loss: 0.3798993229866028

step: 60, loss: 0.2227325737476349

step: 70, loss: 0.20521916449069977

step: 80, loss: 0.105820432305336

step: 90, loss: 0.19560301303863525

step: 100, loss: 0.1545707732439041

step: 110, loss: 0.19842161238193512

step: 120, loss: 0.39214619994163513

step: 130, loss: 0.22692032158374786

step: 140, loss: 0.14939731359481812

step: 150, loss: 0.20017389953136444

step: 160, loss: 0.19623328745365143

step: 170, loss: 0.15158611536026

step: 180, loss: 0.11652106791734695

```

step: 190, loss: 0.20102299749851227
step: 200, loss: 0.11259198933839798
step: 210, loss: 0.018986262381076813
step: 220, loss: 0.08419016003608704
step: 230, loss: 0.21794740855693817
step: 240, loss: 0.041142962872982025
step: 250, loss: 0.05590023100376129
step: 260, loss: 0.3458285331726074
step: 270, loss: 0.33452603220939636
step: 280, loss: 0.04522467404603958
step: 290, loss: 0.06635979562997818
step: 300, loss: 0.19878576695919037
step: 310, loss: 0.22761908173561096
step: 320, loss: 0.3146844804286957
step: 330, loss: 0.10536574572324753
step: 340, loss: 0.3581877648830414
step: 350, loss: 0.19319921731948853
step: 360, loss: 0.030142955482006073
step: 370, loss: 0.03950817883014679
step: 380, loss: 0.2273423969745636
step: 390, loss: 0.10199730843305588
step: 400, loss: 0.04267338663339615
step: 410, loss: 0.2683485448360443
step: 420, loss: 0.02021108567714691
step: 430, loss: 0.030020756646990776
step: 440, loss: 0.045066121965646744
step: 450, loss: 0.23802755773067474
step: 460, loss: 0.0408393070101738
step: 470, loss: 0.055891409516334534
step: 480, loss: 0.09586033970117569
step: 490, loss: 0.15885592997074127
step: 500, loss: 0.09679906815290451
step: 510, loss: 0.19624102115631104
step: 520, loss: 0.09753461182117462
step: 530, loss: 0.0509609617292881
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9181818181818181, f1=0.9100045682960255,
best_f1=0.9100045682960255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08216872066259384

```

step: 10, loss: 0.10407128930091858
step: 20, loss: 0.0444609634578228
step: 30, loss: 0.04560723155736923
step: 40, loss: 0.15325501561164856
step: 50, loss: 0.0884077250957489
step: 60, loss: 0.05525137111544609
step: 70, loss: 0.03652910888195038
step: 80, loss: 0.03391186520457268
step: 90, loss: 0.03220114856958389
step: 100, loss: 0.1456587016582489
step: 110, loss: 0.0202960092574358
step: 120, loss: 0.11420410871505737
step: 130, loss: 0.04793522506952286
step: 140, loss: 0.10709229856729507
step: 150, loss: 0.12325486540794373
step: 160, loss: 0.03841632232069969
step: 170, loss: 0.06005676090717316
step: 180, loss: 0.17753253877162933
step: 190, loss: 0.009008681401610374
step: 200, loss: 0.2482953518629074
step: 210, loss: 0.04763041064143181
step: 220, loss: 0.005925398785620928
step: 230, loss: 0.17700067162513733
step: 240, loss: 0.014624780975282192
step: 250, loss: 0.08231694996356964
step: 260, loss: 0.07207413017749786
step: 270, loss: 0.032007958739995956
step: 280, loss: 0.09062787145376205
step: 290, loss: 0.039600152522325516
step: 300, loss: 0.07312898337841034
step: 310, loss: 0.019414415583014488
step: 320, loss: 0.05935566872358322
step: 330, loss: 0.07440978288650513
step: 340, loss: 0.0953124389052391
step: 350, loss: 0.004342540167272091
step: 360, loss: 0.05129224434494972
step: 370, loss: 0.017414642497897148
step: 380, loss: 0.17271126806735992
step: 390, loss: 0.011242740787565708
step: 400, loss: 0.20970435440540314
step: 410, loss: 0.19538481533527374
step: 420, loss: 0.018675556406378746
step: 430, loss: 0.041912540793418884
step: 440, loss: 0.016427230089902878
step: 450, loss: 0.051229462027549744
step: 460, loss: 0.07360034435987473
step: 470, loss: 0.10207147896289825
step: 480, loss: 0.015447210520505905

```

step: 490, loss: 0.18423186242580414
step: 500, loss: 0.019812053069472313
step: 510, loss: 0.17371107637882233
step: 520, loss: 0.4413681626319885
step: 530, loss: 0.15628862380981445
/usr/local/lib/python3.7/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=0.9158708503865394, f1=0.9124886052871468,
best_f1=0.9100045682960255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1281951665878296
step: 10, loss: 0.20190073549747467
step: 20, loss: 0.02211545966565609
step: 30, loss: 0.06070111691951752
step: 40, loss: 0.0954495444893837
step: 50, loss: 0.03385042026638985
step: 60, loss: 0.012383269146084785
step: 70, loss: 0.06137944012880325
step: 80, loss: 0.07517939060926437
step: 90, loss: 0.026372859254479408
step: 100, loss: 0.1015608087182045
step: 110, loss: 0.05124542489647865
step: 120, loss: 0.11945945769548416
step: 130, loss: 0.04059898853302002
step: 140, loss: 0.024837683886289597
step: 150, loss: 0.035854946821928024
step: 160, loss: 0.22911812365055084
step: 170, loss: 0.03175603970885277
step: 180, loss: 0.03350360319018364
step: 190, loss: 0.008206138387322426
step: 200, loss: 0.03466416895389557
step: 210, loss: 0.04382472485303879
step: 220, loss: 0.12132017314434052
step: 230, loss: 0.04190938174724579
step: 240, loss: 0.11375036090612411
step: 250, loss: 0.2167806476354599
step: 260, loss: 0.1291973739862442
step: 270, loss: 0.014301649294793606
step: 280, loss: 0.026718558743596077
step: 290, loss: 0.10671976953744888
step: 300, loss: 0.23884230852127075
step: 310, loss: 0.06937729567289352
step: 320, loss: 0.022478953003883362
step: 330, loss: 0.03274247422814369
step: 340, loss: 0.017731381580233574

```

```

step: 350, loss: 0.08451211452484131
step: 360, loss: 0.00448169931769371
step: 370, loss: 0.028700657188892365
step: 380, loss: 0.019112661480903625
step: 390, loss: 0.006007126532495022
step: 400, loss: 0.15219531953334808
step: 410, loss: 0.02075961045920849
step: 420, loss: 0.007852469570934772
step: 430, loss: 0.0920446589589119
step: 440, loss: 0.20476889610290527
step: 450, loss: 0.023537492379546165
step: 460, loss: 0.1186978742480278
step: 470, loss: 0.025063149631023407
step: 480, loss: 0.10341240465641022
step: 490, loss: 0.013615871779620647
step: 500, loss: 0.014540459029376507
step: 510, loss: 0.025611624121665955
step: 520, loss: 0.03545744717121124
step: 530, loss: 0.060473810881376266
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9363295880149813, f1=0.9102990033222591,
best_f1=0.9102990033222591
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017740746960043907
step: 10, loss: 0.024152478203177452
step: 20, loss: 0.09353101253509521
step: 30, loss: 0.13517743349075317
step: 40, loss: 0.013374621979892254
step: 50, loss: 0.003397943452000618
step: 60, loss: 0.03342008218169212
step: 70, loss: 0.03819878771901131
step: 80, loss: 0.11817427724599838
step: 90, loss: 0.10363473743200302
step: 100, loss: 0.0028715450316667557
step: 110, loss: 0.3034752607345581
step: 120, loss: 0.040742769837379456
step: 130, loss: 0.10061364620923996
step: 140, loss: 0.025917144492268562
step: 150, loss: 0.009253119118511677
step: 160, loss: 0.02056056633591652

```

```

step: 170, loss: 0.016012420877814293
step: 180, loss: 0.015098053961992264
step: 190, loss: 0.016438771039247513
step: 200, loss: 0.004406225401908159
step: 210, loss: 0.0011433707550168037
step: 220, loss: 0.012162022292613983
step: 230, loss: 0.013290273025631905
step: 240, loss: 0.0070624700747430325
step: 250, loss: 0.15686240792274475
step: 260, loss: 0.01599464751780033
step: 270, loss: 0.10272775590419769
step: 280, loss: 0.0026139444671571255
step: 290, loss: 0.030392002314329147
step: 300, loss: 0.0078567024320364
step: 310, loss: 0.040546588599681854
step: 320, loss: 0.0326363667845726
step: 330, loss: 0.009614333510398865
step: 340, loss: 0.005727630574256182
step: 350, loss: 0.005338788032531738
step: 360, loss: 0.034018177539110184
step: 370, loss: 0.006155801471322775
step: 380, loss: 0.006077714264392853
step: 390, loss: 0.0002451346954330802
step: 400, loss: 0.030658552423119545
step: 410, loss: 0.04773630201816559
step: 420, loss: 0.02766355127096176
step: 430, loss: 0.007065534126013517
step: 440, loss: 0.0040089706890285015
step: 450, loss: 0.11489245295524597
step: 460, loss: 0.024882089346647263
step: 470, loss: 0.002683642553165555
step: 480, loss: 0.06647050380706787
step: 490, loss: 0.006339612882584333
step: 500, loss: 0.05354165658354759
step: 510, loss: 0.12012383341789246
step: 520, loss: 0.013777555897831917
step: 530, loss: 0.1304786652326584
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9418334108887855, f1=0.9270346117867166,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003112774807959795
step: 10, loss: 0.04678045213222504
step: 20, loss: 0.010444516316056252
step: 30, loss: 0.03653524070978165
step: 40, loss: 0.012537037022411823
step: 50, loss: 0.11854472756385803
step: 60, loss: 0.02803504467010498
step: 70, loss: 0.012632050551474094
step: 80, loss: 0.0004783446784131229
step: 90, loss: 0.03087322600185871
step: 100, loss: 0.01669938676059246
step: 110, loss: 0.022502539679408073
step: 120, loss: 0.012632324360311031
step: 130, loss: 0.0040377541445195675
step: 140, loss: 0.0030437635723501444
step: 150, loss: 0.019053051248192787
step: 160, loss: 0.061746995896101
step: 170, loss: 0.020911769941449165
step: 180, loss: 0.022249441593885422
step: 190, loss: 0.0018891753861680627
step: 200, loss: 0.0033137453719973564
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 210, loss: 0.05278677120804787
step: 220, loss: 0.01649647206068039
step: 230, loss: 0.01480130199342966
step: 240, loss: 0.029989516362547874
step: 250, loss: 0.1408003568649292
step: 260, loss: 0.0013616118812933564
step: 270, loss: 0.004064307082444429
step: 280, loss: 0.007299345452338457
step: 290, loss: 0.0005937184323556721
step: 300, loss: 0.19465281069278717
step: 310, loss: 0.03904904052615166
step: 320, loss: 0.15952955186367035
step: 330, loss: 0.0245680995285511
step: 340, loss: 0.0047199479304254055
step: 350, loss: 0.004466185346245766
step: 360, loss: 0.0007034278241917491
step: 370, loss: 0.022570542991161346
step: 380, loss: 0.0008143170853145421
step: 390, loss: 0.010663309134542942
step: 400, loss: 0.022054653614759445
step: 410, loss: 0.011475787498056889
step: 420, loss: 0.2553100287914276
step: 430, loss: 0.11252263188362122
step: 440, loss: 0.0018837377429008484
step: 450, loss: 0.020967863500118256
```

```

step: 460, loss: 0.09108888357877731
step: 470, loss: 0.008361663669347763
step: 480, loss: 0.020784573629498482
step: 490, loss: 0.005315558984875679
step: 500, loss: 0.0497484989464283
step: 510, loss: 0.002075841184705496
step: 520, loss: 0.021061161532998085
step: 530, loss: 0.06867124140262604
/usr/local/lib/python3.7/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=0.9371271225332722, f1=0.9160092807424594,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032231705263257027
step: 10, loss: 0.0052661900408566
step: 20, loss: 0.001944153686054051
step: 30, loss: 0.009063954465091228
step: 40, loss: 0.002740789670497179
step: 50, loss: 0.002777877962216735
step: 60, loss: 0.01143517903983593
step: 70, loss: 0.0019823124166578054
step: 80, loss: 0.005612978246062994
step: 90, loss: 0.07263123989105225
step: 100, loss: 0.1506366729736328
step: 110, loss: 0.007773080375045538
step: 120, loss: 0.0759362131357193
step: 130, loss: 0.014120414853096008
step: 140, loss: 0.003606649348512292
step: 150, loss: 0.00430792523548007
step: 160, loss: 0.01358600053936243
step: 170, loss: 0.005140029825270176
step: 180, loss: 0.0012397004757076502
step: 190, loss: 0.09376617521047592
step: 200, loss: 0.010960179381072521
step: 210, loss: 0.0016084349481388927
step: 220, loss: 0.010106639005243778
step: 230, loss: 0.003992223180830479
step: 240, loss: 0.0033821638207882643
step: 250, loss: 0.09563019126653671
step: 260, loss: 0.030062740668654442
step: 270, loss: 0.043088093400001526
step: 280, loss: 0.003488582791760564
step: 290, loss: 0.0046455408446490765
step: 300, loss: 0.014161588624119759
step: 310, loss: 0.011419082060456276

```

```

step: 320, loss: 0.0008239186136052012
step: 330, loss: 0.04355624318122864
step: 340, loss: 0.0004006222588941455
step: 350, loss: 0.0009523734916001558
step: 360, loss: 0.0035273958928883076
step: 370, loss: 0.003804554929956794
step: 380, loss: 0.0001152541080955416
step: 390, loss: 0.004052366595715284
step: 400, loss: 0.029163029044866562
step: 410, loss: 0.015263229608535767
step: 420, loss: 0.004885219503194094
step: 430, loss: 0.05097469314932823
step: 440, loss: 0.0036637485027313232
step: 450, loss: 0.145896315574646
step: 460, loss: 0.006567580159753561
step: 470, loss: 0.0065543800592422485
step: 480, loss: 0.003086814656853676
step: 490, loss: 0.004923861473798752
step: 500, loss: 0.0035533271729946136
step: 510, loss: 0.0007419213652610779
step: 520, loss: 0.0015977936564013362
step: 530, loss: 0.0015265787951648235
/usr/local/lib/python3.7/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=0.9336426914153132, f1=0.9177570093457944,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019972561858594418
step: 10, loss: 0.0010174672352150083
step: 20, loss: 0.012327433563768864
step: 30, loss: 0.05608770623803139
step: 40, loss: 0.0008528506732545793
step: 50, loss: 0.0070818024687469006
step: 60, loss: 0.0011145583121106029
step: 70, loss: 0.0003988793177995831
step: 80, loss: 0.013931389898061752
step: 90, loss: 0.0006839275592938066
step: 100, loss: 0.03264468535780907
step: 110, loss: 0.0013856159057468176
step: 120, loss: 0.01447706762701273
step: 130, loss: 0.0025125322863459587
step: 140, loss: 0.0017170746577903628
step: 150, loss: 0.0016028754180297256
step: 160, loss: 0.0009846563916653395
step: 170, loss: 0.0027253341395407915

```

```

step: 180, loss: 0.014016491360962391
step: 190, loss: 0.03538395091891289
step: 200, loss: 0.0004349146329332143
step: 210, loss: 0.0016762695740908384
step: 220, loss: 0.0003111833066213876
step: 230, loss: 0.005899286363273859
step: 240, loss: 0.007451179437339306
step: 250, loss: 0.0014726620865985751
step: 260, loss: 0.0006727228756062686
step: 270, loss: 0.00028299426776356995
step: 280, loss: 0.00050592195475474
step: 290, loss: 0.0014143413864076138
step: 300, loss: 0.0002561870205681771
step: 310, loss: 0.0019483452197164297
step: 320, loss: 0.04228222742676735
step: 330, loss: 0.07709917426109314
step: 340, loss: 0.002070090500637889
step: 350, loss: 0.004841519985347986
step: 360, loss: 0.05371318385004997
step: 370, loss: 0.003983228001743555
step: 380, loss: 0.0021381909027695656
step: 390, loss: 0.01677568629384041
step: 400, loss: 0.011851407587528229
step: 410, loss: 0.00023409645655192435
step: 420, loss: 0.02535650134086609
step: 430, loss: 0.007650649640709162
step: 440, loss: 0.0005736211896874011
step: 450, loss: 0.0061559248715639114
step: 460, loss: 0.0027885681483894587
step: 470, loss: 0.13499052822589874
step: 480, loss: 0.008310130797326565
step: 490, loss: 0.0012001382419839501
step: 500, loss: 0.0007449891418218613
step: 510, loss: 0.0029760291799902916
step: 520, loss: 0.0007221864070743322
step: 530, loss: 0.0015609099064022303
/usr/local/lib/python3.7/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=0.9412861136999067, f1=0.91828058573453,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008670588140375912
step: 10, loss: 0.0029360430780798197
step: 20, loss: 0.003495139069855213
step: 30, loss: 0.04880039766430855

```

step: 40, loss: 0.00043486899812705815
step: 50, loss: 0.0015068453503772616
step: 60, loss: 0.0032000758219510317
step: 70, loss: 0.1198335662484169
step: 80, loss: 0.027225850149989128
step: 90, loss: 0.002965047024190426
step: 100, loss: 0.007277938537299633
step: 110, loss: 0.0004200826515443623
step: 120, loss: 0.00023601812426932156
step: 130, loss: 0.0012922120513394475
step: 140, loss: 0.0014485253486782312
step: 150, loss: 0.00015217013424262404
step: 160, loss: 0.0068990690633654594
step: 170, loss: 0.10897748917341232
step: 180, loss: 0.0033542602322995663
step: 190, loss: 0.006716785952448845
step: 200, loss: 0.025929193943738937
step: 210, loss: 0.022712426260113716
step: 220, loss: 0.05580299720168114
step: 230, loss: 0.020299164578318596
step: 240, loss: 0.004297539591789246
step: 250, loss: 0.0004684797895606607
step: 260, loss: 0.0001974530896404758
step: 270, loss: 0.01809442602097988
step: 280, loss: 0.00034963051439262927
step: 290, loss: 0.0017137560062110424
step: 300, loss: 0.00017939215467777103
step: 310, loss: 0.0005525902379304171
step: 320, loss: 0.00043977960012853146
step: 330, loss: 0.00020773608412127942
step: 340, loss: 0.0054682050831615925
step: 350, loss: 7.808498048689216e-05
step: 360, loss: 0.009248107671737671
step: 370, loss: 0.0006262484821490943
step: 380, loss: 0.001749142655171454
step: 390, loss: 0.00045138399582356215
step: 400, loss: 0.0009482354507781565
step: 410, loss: 0.0016647193115204573
step: 420, loss: 0.000579641608055681
step: 430, loss: 0.0019841522444039583
step: 440, loss: 0.021275555714964867
step: 450, loss: 0.0023179673589766026
step: 460, loss: 0.0029674572870135307
step: 470, loss: 0.02874019555747509
step: 480, loss: 0.01507902517914772
step: 490, loss: 0.000966049381531775
step: 500, loss: 0.06006579101085663
step: 510, loss: 0.016779735684394836

```

step: 520, loss: 0.000248784723225981
step: 530, loss: 0.0010594648774713278
/usr/local/lib/python3.7/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=0.9400921658986175, f1=0.9306839186691312,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005400690715759993
step: 10, loss: 0.00455357413738966
step: 20, loss: 0.14729678630828857
step: 30, loss: 0.06636299937963486
step: 40, loss: 0.02002984844148159
step: 50, loss: 0.0013357005082070827
step: 60, loss: 0.001295054447837174
step: 70, loss: 0.0015974657144397497
step: 80, loss: 0.002402914920821786
step: 90, loss: 0.06168947368860245
step: 100, loss: 0.0018179869512096047
step: 110, loss: 0.016451887786388397
step: 120, loss: 0.0028222985565662384
step: 130, loss: 0.0002632595715112984
step: 140, loss: 0.0008366734837181866
step: 150, loss: 0.017088737338781357
step: 160, loss: 0.0009003740851767361
step: 170, loss: 0.016035055741667747
step: 180, loss: 0.0005946720484644175
step: 190, loss: 0.00016004905046429485
step: 200, loss: 9.730477177072316e-05
step: 210, loss: 0.07331323623657227
step: 220, loss: 0.00019526334654074162
step: 230, loss: 0.000994460890069604
step: 240, loss: 0.0013114181347191334
step: 250, loss: 0.006756719667464495
step: 260, loss: 0.03254237771034241
step: 270, loss: 0.01551765762269497
step: 280, loss: 0.0007394292042590678
step: 290, loss: 0.0006594533915631473
step: 300, loss: 0.0013591154711320996
step: 310, loss: 0.00013994322216603905
step: 320, loss: 0.00021131498215254396
step: 330, loss: 0.00028128564008511603
step: 340, loss: 0.001033903332427144
step: 350, loss: 0.005330143496394157
step: 360, loss: 0.0014489793684333563
step: 370, loss: 0.0006859877030365169

```

```

step: 380, loss: 0.000948809552937746
step: 390, loss: 0.000632519309874624
step: 400, loss: 0.0016774858813732862
step: 410, loss: 0.0011411579325795174
step: 420, loss: 0.0003996138402726501
step: 430, loss: 0.16378408670425415
step: 440, loss: 0.0004505810502450913
step: 450, loss: 0.0038017896004021168
step: 460, loss: 0.018794726580381393
step: 470, loss: 0.00022169531439431012
step: 480, loss: 0.00011736510350601748
step: 490, loss: 0.0017313417047262192
step: 500, loss: 0.0047581931576132774
step: 510, loss: 0.002745227422565222
step: 520, loss: 0.03240828588604927
step: 530, loss: 0.01672876439988613
/usr/local/lib/python3.7/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=0.9407372841810545, f1=0.9258049463369109,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0043505108915269375
step: 10, loss: 0.0004302221059333533
step: 20, loss: 0.00020156337996013463
step: 30, loss: 0.003431012388318777
step: 40, loss: 0.0006283809198066592
step: 50, loss: 0.0006900609587319195
step: 60, loss: 0.0006405394524335861
step: 70, loss: 0.000403459562221542
step: 80, loss: 0.0002346794935874641
step: 90, loss: 0.0002884833374992013
step: 100, loss: 0.055600959807634354
step: 110, loss: 0.05352189391851425
step: 120, loss: 0.0016055564628913999
step: 130, loss: 0.0007180828251875937
step: 140, loss: 0.0042383503168821335
step: 150, loss: 0.001070067286491394
step: 160, loss: 0.04516424238681793
step: 170, loss: 0.0006642877706326544
step: 180, loss: 0.012463822960853577
step: 190, loss: 0.0033127893693745136
step: 200, loss: 0.00015709105355199426
step: 210, loss: 0.004282026551663876
step: 220, loss: 0.0005239093443378806
step: 230, loss: 0.00027042353758588433

```

```

step: 240, loss: 0.0002982448786497116
step: 250, loss: 0.00020043857512064278
step: 260, loss: 0.005187671165913343
step: 270, loss: 0.0008195448899641633
step: 280, loss: 0.00620149215683341
step: 290, loss: 0.0007553101750090718
step: 300, loss: 0.0007445216178894043
step: 310, loss: 0.09090220928192139
step: 320, loss: 0.001942290342412889
step: 330, loss: 0.021671701222658157
step: 340, loss: 0.00013287724868860096
step: 350, loss: 0.0003977588494308293
step: 360, loss: 0.00015531932876911014
step: 370, loss: 0.0001744187029544264
step: 380, loss: 0.002327019115909934
step: 390, loss: 0.00012330089521128684
step: 400, loss: 0.000220409594476223
step: 410, loss: 0.00014432996977120638
step: 420, loss: 0.0032450095750391483
step: 430, loss: 0.0016158074140548706
step: 440, loss: 0.0006097028963267803
step: 450, loss: 0.03535338118672371
step: 460, loss: 0.001220375532284379
step: 470, loss: 0.010670822113752365
step: 480, loss: 0.00022483928478322923
step: 490, loss: 6.896322884131223e-05
step: 500, loss: 0.0004373928240966052
step: 510, loss: 3.0318606150103733e-05
step: 520, loss: 5.9548292483668774e-05
step: 530, loss: 0.0008881750982254744
/usr/local/lib/python3.7/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=0.9394650398873767, f1=0.9273323956868261,
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.133209975203499e-05
step: 10, loss: 7.148718577809632e-05
step: 20, loss: 0.01534039806574583
step: 30, loss: 3.91099565604236e-05
step: 40, loss: 0.0001856351736932993
step: 50, loss: 3.4937278542201966e-05
step: 60, loss: 6.857769767520949e-05
step: 70, loss: 0.0014081698609516025
step: 80, loss: 0.0048409332521259785
step: 90, loss: 0.0009406769531778991

```



```
step: 100, loss: 0.0011187989730387926
step: 110, loss: 0.00011050880857510492
step: 120, loss: 0.0004868104588240385
step: 130, loss: 5.011531902709976e-05
step: 140, loss: 6.625335663557053e-05
step: 150, loss: 0.0011347298277541995
step: 160, loss: 6.849091005278751e-05
step: 170, loss: 9.782981942407787e-05
step: 180, loss: 4.431097113410942e-05
step: 190, loss: 0.00015937944408506155
step: 200, loss: 1.5813600839464925e-05
step: 210, loss: 0.0001200932965730317
step: 220, loss: 0.0012752312468364835
step: 230, loss: 2.4783206754364073e-05
step: 240, loss: 0.002875553211197257
step: 250, loss: 8.480856922687963e-05
step: 260, loss: 0.000505251984577626
step: 270, loss: 0.004018457140773535
step: 280, loss: 5.067258098279126e-05
step: 290, loss: 4.1906347178155556e-05
step: 300, loss: 0.0008242235635407269
step: 310, loss: 0.000680859200656414
step: 320, loss: 0.004668467212468386
step: 330, loss: 3.838686097878963e-05
step: 340, loss: 0.0013241316191852093
step: 350, loss: 4.3599593482213095e-05
step: 360, loss: 5.1582555897766724e-05
step: 370, loss: 0.0005813354509882629
step: 380, loss: 0.004166029393672943
step: 390, loss: 0.00668302970007062
step: 400, loss: 2.797221895889379e-05
step: 410, loss: 0.00029709129012189806
step: 420, loss: 0.05573005974292755
step: 430, loss: 0.00019820783927571028
step: 440, loss: 2.4221211788244545e-05
step: 450, loss: 0.0002484689757693559
step: 460, loss: 0.0006522481562569737
step: 470, loss: 0.0002157177368644625
step: 480, loss: 1.8153046767110936e-05
step: 490, loss: 3.0185921787051484e-05
step: 500, loss: 4.236742097418755e-05
step: 510, loss: 1.9464188881102018e-05
step: 520, loss: 7.462058420060202e-05
step: 530, loss: 3.694889892358333e-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=0.9407407407407408, f1=0.9290681502086231,
```

```
best_f1=0.9270346117867166
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.811357212020084e-05
step: 10, loss: 7.56334702600725e-05
step: 20, loss: 6.285551353357732e-05
step: 30, loss: 1.1838787941087503e-05
step: 40, loss: 3.549527173163369e-05
step: 50, loss: 0.00026174724916927516
step: 60, loss: 0.00011702445044647902
step: 70, loss: 0.08110152184963226
step: 80, loss: 4.428827378433198e-05
step: 90, loss: 0.00039002051926217973
step: 100, loss: 0.05332622677087784
step: 110, loss: 0.007422544993460178
step: 120, loss: 0.00011817103222711012
step: 130, loss: 0.00030647983658127487
step: 140, loss: 0.00021949890651740134
step: 150, loss: 0.0006787551683373749
step: 160, loss: 0.002636122517287731
step: 170, loss: 7.788729999447241e-05
step: 180, loss: 0.000408309482736513
step: 190, loss: 6.945132918190211e-05
step: 200, loss: 8.320189226651564e-05
step: 210, loss: 0.0009252433083020151
step: 220, loss: 4.295519465813413e-05
step: 230, loss: 4.591992183122784e-05
step: 240, loss: 0.0022241142578423023
step: 250, loss: 0.00015492341481149197
step: 260, loss: 3.556919182301499e-05
step: 270, loss: 2.839993976522237e-05
step: 280, loss: 0.014067750424146652
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 290, loss: 5.7147652114508674e-05
step: 300, loss: 0.00020694313570857048
step: 310, loss: 0.001603981712833047
step: 320, loss: 0.03436854109168053
step: 330, loss: 0.0002148322091670707
step: 340, loss: 0.02788003720343113
step: 350, loss: 9.074129047803581e-05
step: 360, loss: 0.0006195244495756924
step: 370, loss: 1.2155404874647502e-05
step: 380, loss: 0.0004864012007601559
step: 390, loss: 0.007850281894207
step: 400, loss: 2.5182134777423926e-05
step: 410, loss: 0.0027814649511128664
step: 420, loss: 0.002016404876485467
```

```

step: 430, loss: 0.0008502421551384032
step: 440, loss: 0.0025340032298117876
step: 450, loss: 0.01020789984613657
step: 460, loss: 0.0006560960318893194
step: 470, loss: 0.0023703293409198523
step: 480, loss: 0.0015814048238098621
step: 490, loss: 0.0014464394189417362
step: 500, loss: 4.351206007413566e-05
step: 510, loss: 0.029697036370635033
step: 520, loss: 0.010098441503942013
step: 530, loss: 0.0002945071319118142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.94547134935305, f1=0.9306839186691312,
best_f1=0.9306839186691312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.374403164954856e-05
step: 10, loss: 0.039856646209955215
step: 20, loss: 9.57817246671766e-05
step: 30, loss: 2.925258013419807e-05
step: 40, loss: 0.022888634353876114
step: 50, loss: 0.00011953315697610378
step: 60, loss: 0.0002076198870781809
step: 70, loss: 0.0004918453050777316
step: 80, loss: 0.012960166670382023
step: 90, loss: 3.55513438989874e-05
step: 100, loss: 0.00022909634571988136
step: 110, loss: 0.00013480168126989156
step: 120, loss: 7.114997424650937e-05
step: 130, loss: 0.00034575717290863395
step: 140, loss: 0.00017930175818037242
step: 150, loss: 0.00013808876974508166
step: 160, loss: 0.0009963734773918986
step: 170, loss: 0.00015303936379496008
step: 180, loss: 4.3350377382012084e-05
step: 190, loss: 5.121979120303877e-05
step: 200, loss: 3.531498441589065e-05
step: 210, loss: 0.0009525195928290486
step: 220, loss: 1.9929577319999225e-05
step: 230, loss: 0.00023816524480935186
step: 240, loss: 0.0007758997962810099

```

```

step: 250, loss: 4.1781971958698705e-05
step: 260, loss: 0.0001537479110993445
step: 270, loss: 6.729971210006624e-05
step: 280, loss: 0.05733225494623184
step: 290, loss: 1.3045742889516987e-05
step: 300, loss: 0.00010849712998606265
step: 310, loss: 8.869838347891346e-05
step: 320, loss: 1.963920658454299e-05
step: 330, loss: 0.040905941277742386
step: 340, loss: 0.00013885532098356634
step: 350, loss: 6.569417746504769e-05
step: 360, loss: 0.05823296681046486
step: 370, loss: 0.011759252287447453
step: 380, loss: 0.0011915276991203427
step: 390, loss: 0.0003266696003265679
step: 400, loss: 0.0002865652786567807
step: 410, loss: 2.5446541258133948e-05
step: 420, loss: 2.353172021685168e-05
step: 430, loss: 3.5636847314890474e-05
step: 440, loss: 2.032441807386931e-05
step: 450, loss: 0.00014391900913324207
step: 460, loss: 9.712758037494496e-05
step: 470, loss: 0.0034952780697494745
step: 480, loss: 2.0047389625688083e-05
step: 490, loss: 1.3991895684739575e-05
step: 500, loss: 3.634481618064456e-05
step: 510, loss: 5.5500462622148916e-05
step: 520, loss: 3.1880779715720564e-05
step: 530, loss: 2.814567051245831e-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=0.9387379087977891, f1=0.9217954650624711,
best_f1=0.9306839186691312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.6925812310073525e-05
step: 10, loss: 6.149403634481132e-05
step: 20, loss: 1.4166998880682513e-05
step: 30, loss: 0.00023499118105974048
step: 40, loss: 2.7109921575174667e-05
step: 50, loss: 0.021635612472891808
step: 60, loss: 0.0005497133824974298
step: 70, loss: 3.352300700498745e-05
step: 80, loss: 4.6916746214265004e-05
step: 90, loss: 1.3060603123449255e-05
step: 100, loss: 2.7177922675036825e-05

```

```
step: 110, loss: 3.9940794522408396e-05
step: 120, loss: 2.398540345893707e-05
step: 130, loss: 1.2859525668318383e-05
step: 140, loss: 0.0001410551049048081
step: 150, loss: 3.310311876703054e-05
step: 160, loss: 3.117047162959352e-05
step: 170, loss: 2.160220901714638e-05
step: 180, loss: 2.7560665330383927e-05
step: 190, loss: 3.282114994362928e-05
step: 200, loss: 3.548320091795176e-05
step: 210, loss: 9.563953790348023e-05
step: 220, loss: 0.0001966948912013322
step: 230, loss: 1.2840913768741302e-05
step: 240, loss: 1.528052416688297e-05
step: 250, loss: 2.3862663510954008e-05
step: 260, loss: 2.404807491984684e-05
step: 270, loss: 3.200861465302296e-05
step: 280, loss: 1.2855703062086832e-05
step: 290, loss: 1.5068262655404396e-05
step: 300, loss: 0.00011651866952888668
step: 310, loss: 4.0981794882100075e-05
step: 320, loss: 1.252054516953649e-05
step: 330, loss: 0.0002471235056873411
step: 340, loss: 4.193964559817687e-05
step: 350, loss: 1.960178451554384e-05
step: 360, loss: 1.9299872292322107e-05
step: 370, loss: 2.5072022253880277e-05
step: 380, loss: 0.0013312924420461059
step: 390, loss: 9.037472409545444e-06
step: 400, loss: 2.8232290787855163e-05
step: 410, loss: 0.00019932363647967577
step: 420, loss: 0.0008958350517787039
step: 430, loss: 0.0001278077543247491
step: 440, loss: 0.0004166796861682087
step: 450, loss: 1.546321072964929e-05
step: 460, loss: 7.85543816164136e-05
step: 470, loss: 2.0011686501675285e-05
step: 480, loss: 1.558625081088394e-05
step: 490, loss: 1.3787116586172488e-05
step: 500, loss: 2.406796374998521e-05
step: 510, loss: 0.00019188941223546863
step: 520, loss: 1.6568941646255553e-05
step: 530, loss: 5.0008056859951466e-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=0.9406858202038925, f1=0.9256505576208177,
best_f1=0.9306839186691312
```

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

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

```
step: 0, loss: 3.327985541545786e-05  
step: 10, loss: 6.0081525589339435e-05  
step: 20, loss: 3.4567925467854366e-05  
step: 30, loss: 9.652680455474183e-05  
step: 40, loss: 0.0007844914798624814  
step: 50, loss: 0.0005304288351908326  
step: 60, loss: 2.0420986402314156e-05  
step: 70, loss: 1.6778165445430204e-05  
step: 80, loss: 1.3481582755048294e-05  
step: 90, loss: 0.00012454346870072186  
step: 100, loss: 7.200939307949739e-06  
step: 110, loss: 1.899455310194753e-05  
step: 120, loss: 1.5347852240665816e-05  
step: 130, loss: 0.00023172864166554064  
step: 140, loss: 1.215172596857883e-05  
step: 150, loss: 1.2002759831375442e-05  
step: 160, loss: 2.3484681150875986e-05  
step: 170, loss: 1.5049585272208788e-05  
step: 180, loss: 1.356733264401555e-05  
step: 190, loss: 9.93942521745339e-05  
step: 200, loss: 1.1421572708059102e-05  
step: 210, loss: 8.88739014044404e-05  
step: 220, loss: 3.423505768296309e-05  
step: 230, loss: 2.047362067969516e-05  
step: 240, loss: 1.2256004993105307e-05  
step: 250, loss: 0.00022391066886484623  
step: 260, loss: 1.2457204320526216e-05  
step: 270, loss: 2.0607314581866376e-05  
step: 280, loss: 2.679432145669125e-05  
step: 290, loss: 0.0005152954836376011  
step: 300, loss: 2.3001879526418634e-05  
step: 310, loss: 5.8000463468488306e-05  
step: 320, loss: 1.473320480727125e-05  
step: 330, loss: 0.00014378408377524465  
step: 340, loss: 2.2443909983849153e-05  
step: 350, loss: 8.68431743583642e-05  
step: 360, loss: 2.8464915885706432e-05  
step: 370, loss: 3.2877182093216106e-05  
step: 380, loss: 1.711739423626568e-05  
step: 390, loss: 0.00011685894423862919  
step: 400, loss: 0.00843277107924223  
step: 410, loss: 1.902017174870707e-05  
step: 420, loss: 1.1868602996401023e-05  
step: 430, loss: 8.832581443130039e-06  
step: 440, loss: 0.00012236012844368815
```

```

step: 450, loss: 0.0006368214380927384
step: 460, loss: 0.000134775647893548
step: 470, loss: 1.1239076229685452e-05
step: 480, loss: 0.0037566591054201126
step: 490, loss: 1.3015968761465047e-05
step: 500, loss: 4.487911064643413e-05
step: 510, loss: 2.351096372876782e-05
step: 520, loss: 1.9322375010233372e-05
step: 530, loss: 1.0330101758881938e-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=0.9411219286045434, f1=0.9253592953175707,
best_f1=0.9306839186691312
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:35, 160.91it/s]
load_f1 = 0.9414498141263941
real_f1 = 0.9383057090239412
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-

```

```
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:30, 143.94it/s]
```

4 DITTO

4.1 DITTO STRUCTURED

4.1.1 Beer - Running the matcher

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

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

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

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


```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled          : True
opt_level        : 02
cast_model_type  : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights   : True
loss_scale       : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.422829806804657
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=0.345679012345679, f1=0.30000000000000004,
best_f1=0.30000000000000004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.4226915240287781
/usr/local/lib/python3.7/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=0.3181818181818182, f1=0.3111111111111111,
best_f1=0.30000000000000004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41993170976638794
/usr/local/lib/python3.7/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=0.34210526315789475, f1=0.35294117647058826,
best_f1=0.30000000000000004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25956565141677856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.423076923076923, f1=0.358974358974359,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28212061524391174
/usr/local/lib/python3.7/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=0.41791044776119407, f1=0.3773584905660377,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2509179711341858
/usr/local/lib/python3.7/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=0.3636363636363636, f1=0.34782608695652173,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5572521686553955
/usr/local/lib/python3.7/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=0.2692307692307693, f1=0.2692307692307693,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.49403059482574463
/usr/local/lib/python3.7/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=0.28, f1=0.2745098039215686, best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36697012186050415
/usr/local/lib/python3.7/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=0.3870967741935484, f1=0.29629629629629634,
best_f1=0.358974358974359
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.43818768858909607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4897959183673469, f1=0.5128205128205129,
best_f1=0.5128205128205129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5211028456687927
/usr/local/lib/python3.7/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=0.43750000000000006, f1=0.42857142857142855,
best_f1=0.5128205128205129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3046155869960785
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5263157894736842, f1=0.5161290322580646,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30589133501052856
/usr/local/lib/python3.7/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=0.4727272727272727, f1=0.5853658536585367,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2553427815437317
/usr/local/lib/python3.7/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=0.4905660377358491, f1=0.5853658536585367,
best_f1=0.5161290322580646
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34741783142089844
/usr/local/lib/python3.7/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=0.4905660377358491, f1=0.5853658536585367,
best_f1=0.5161290322580646
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, 123202.60it/s]
/usr/local/lib/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 = 0.46153846153846156
real_f1 = 0.43750000000000006
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:30, 141.92it/s]

```

4.1.2 DBLP-ACM - Running the matcher

```

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

```

```

--summarize \
--da entry_swap \
--da attr_del

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

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

step: 0, loss: 0.5636671781539917

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

step: 10, loss: 0.44711920619010925

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.5398500561714172
step: 30, loss: 0.2914138436317444
step: 40, loss: 0.3590284287929535
step: 50, loss: 0.5874623656272888
step: 60, loss: 0.480815589427948
step: 70, loss: 0.41552168130874634
step: 80, loss: 0.6294143199920654
step: 90, loss: 0.44097548723220825
step: 100, loss: 0.5174350738525391
step: 110, loss: 0.5943555235862732
step: 120, loss: 0.4370608925819397
step: 130, loss: 0.35224977135658264
step: 140, loss: 0.3821382522583008
step: 150, loss: 0.5593423247337341
step: 160, loss: 0.5098389983177185
step: 170, loss: 0.5057608485221863
step: 180, loss: 0.3403456509113312
step: 190, loss: 0.10818175226449966
step: 200, loss: 0.1743936985731125
step: 210, loss: 0.2014603614807129
step: 220, loss: 0.2701810896396637
step: 230, loss: 0.009849537163972855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8631813125695216, f1=0.8820286659316429,
best_f1=0.8820286659316429
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05433689430356026
step: 10, loss: 0.210844486951828
step: 20, loss: 0.09981407225131989
step: 30, loss: 0.08081388473510742
step: 40, loss: 0.06457076221704483
step: 50, loss: 0.011023938655853271
step: 60, loss: 0.016166800633072853
step: 70, loss: 0.16654027998447418
step: 80, loss: 0.011722440831363201
step: 90, loss: 0.03946695104241371
step: 100, loss: 0.06059879809617996
step: 110, loss: 0.08606406301259995

```

```

step: 120, loss: 0.004168601240962744
step: 130, loss: 0.17201294004917145
step: 140, loss: 0.009984279051423073
step: 150, loss: 0.044787075370550156
step: 160, loss: 0.01437772624194622
step: 170, loss: 0.026120884343981743
step: 180, loss: 0.008202768862247467
step: 190, loss: 0.0596257708966732
step: 200, loss: 0.009162255562841892
step: 210, loss: 0.22084787487983704
step: 220, loss: 0.0530737042427063
step: 230, loss: 0.03470532223582268
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9611542730299667, f1=0.9709821428571428,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02519683726131916
step: 10, loss: 0.013266921043395996
step: 20, loss: 0.012450497597455978
step: 30, loss: 0.001527939923107624
step: 40, loss: 0.046315573155879974
step: 50, loss: 0.005182642489671707
step: 60, loss: 0.021446356549859047
step: 70, loss: 0.006889257580041885
step: 80, loss: 0.0030678242910653353
step: 90, loss: 0.024150514975190163
step: 100, loss: 0.018364770337939262
step: 110, loss: 0.009485527873039246
step: 120, loss: 0.032188281416893005
step: 130, loss: 0.0037102934438735247
step: 140, loss: 0.046992044895887375
step: 150, loss: 0.09248426556587219
step: 160, loss: 0.004998256918042898
step: 170, loss: 0.0015767848817631602
step: 180, loss: 0.040052078664302826
step: 190, loss: 0.07959777861833572
step: 200, loss: 0.07181651890277863
step: 210, loss: 0.005230489186942577
step: 220, loss: 0.025675782933831215
step: 230, loss: 0.028622429817914963

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842696629213483,
best_f1=0.9842696629213483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010147296823561192
step: 10, loss: 0.0035419638734310865
step: 20, loss: 0.06534707546234131
step: 30, loss: 0.0017445062985643744
step: 40, loss: 0.16957612335681915
step: 50, loss: 0.0077243996784091
step: 60, loss: 0.017688974738121033
step: 70, loss: 0.1847972720861435
step: 80, loss: 0.006782884243875742
step: 90, loss: 0.028220057487487793
step: 100, loss: 0.0161898173391819
step: 110, loss: 0.0015684411628171802
step: 120, loss: 0.0014443229883909225
step: 130, loss: 0.011808088980615139
step: 140, loss: 0.0003217694174963981
step: 150, loss: 0.0002962341532111168
step: 160, loss: 0.006052056327462196
step: 170, loss: 0.0801510140299797
step: 180, loss: 0.12716278433799744
step: 190, loss: 0.0009822896681725979
step: 200, loss: 0.0240604467689991
step: 210, loss: 0.004464657977223396
step: 220, loss: 0.004504515323787928
step: 230, loss: 0.002690643072128296
/usr/local/lib/python3.7/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=0.9786276715410572, f1=0.9831271091113611,
best_f1=0.9842696629213483
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01652432046830654
step: 10, loss: 0.003782259998843074
step: 20, loss: 0.08296452462673187
step: 30, loss: 0.0012121874606236815

```

```

step: 40, loss: 0.0049170684069395065
step: 50, loss: 0.012001731432974339
step: 60, loss: 0.034413158893585205
step: 70, loss: 0.05344155803322792
step: 80, loss: 0.03490067273378372
step: 90, loss: 0.029498320072889328
step: 100, loss: 0.0002416380593786016
step: 110, loss: 0.009365974925458431
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 120, loss: 0.05940457060933113
step: 130, loss: 0.0250677652657032
step: 140, loss: 0.08047596365213394
step: 150, loss: 0.15672282874584198
step: 160, loss: 0.03270735964179039
step: 170, loss: 0.08840573579072952
step: 180, loss: 0.011220552958548069
step: 190, loss: 0.011949364095926285
step: 200, loss: 0.015168348327279091
step: 210, loss: 0.002994536655023694
step: 220, loss: 0.01017130445688963
step: 230, loss: 0.011136564426124096
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 5: dev_f1=0.983050847457627, f1=0.9875424688561721,
best_f1=0.9875424688561721
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001296313712373376
step: 10, loss: 0.023783622309565544
step: 20, loss: 0.0013111650478094816
step: 30, loss: 0.021091589704155922
step: 40, loss: 0.0004969686851836741
step: 50, loss: 0.008713659830391407
step: 60, loss: 0.031143661588430405
step: 70, loss: 0.00043924301280640066
step: 80, loss: 0.013029232621192932
step: 90, loss: 0.016282090917229652
step: 100, loss: 0.0008650583913549781
step: 110, loss: 0.11362996697425842
step: 120, loss: 0.0002959483826998621
step: 130, loss: 0.05266730859875679
step: 140, loss: 0.006631330121308565

```

```

step: 150, loss: 0.002339593367651105
step: 160, loss: 0.006471929140388966
step: 170, loss: 0.00045578714343719184
step: 180, loss: 0.00033229784457944334
step: 190, loss: 0.0027514277026057243
step: 200, loss: 0.009115728549659252
step: 210, loss: 0.001321954419836402
step: 220, loss: 0.10745427012443542
step: 230, loss: 0.0004528223071247339
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.984304932735426, f1=0.9784824462061155,
best_f1=0.9784824462061155
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016018958762288094
step: 10, loss: 0.000506450654938817
step: 20, loss: 0.0001506220578448847
step: 30, loss: 8.598714339314029e-05
step: 40, loss: 0.0016296057729050517
step: 50, loss: 0.0002877672668546438
step: 60, loss: 0.00018288390128873289
step: 70, loss: 0.0002951862115878612
step: 80, loss: 0.0038829382974654436
step: 90, loss: 0.00144118198659271
step: 100, loss: 0.001063240459188819
step: 110, loss: 0.0001849825493991375
step: 120, loss: 0.0031573795713484287
step: 130, loss: 0.0001700633147265762
step: 140, loss: 0.00024041591677814722
step: 150, loss: 0.015927033498883247
step: 160, loss: 0.00023017071362119168
step: 170, loss: 0.0023829860147088766
step: 180, loss: 0.004143036901950836
step: 190, loss: 0.0001193142743431963
step: 200, loss: 0.003686124226078391
step: 210, loss: 0.00027161481557413936
step: 220, loss: 0.00014468046720139682
step: 230, loss: 0.003039587987586856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9844097995545658, f1=0.9842342342342343,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022214764612726867
step: 10, loss: 0.002521441550925374
step: 20, loss: 0.001158541301265359
step: 30, loss: 0.0003226103144697845
step: 40, loss: 0.025545889511704445
step: 50, loss: 0.000447397178504616
step: 60, loss: 0.00041365320794284344
step: 70, loss: 0.00015426345635205507
step: 80, loss: 0.00044039799831807613
step: 90, loss: 0.00017898519581649452
step: 100, loss: 0.0016063280636444688
step: 110, loss: 0.008756055496633053
step: 120, loss: 0.00046042934991419315
step: 130, loss: 0.0005618403665721416
step: 140, loss: 0.0003728965821210295
step: 150, loss: 0.02048114873468876
step: 160, loss: 0.0017063500126823783
step: 170, loss: 0.042576227337121964
step: 180, loss: 0.0033891659695655107
step: 190, loss: 0.002116963267326355
step: 200, loss: 9.613288420950994e-05
step: 210, loss: 0.0023527902085334063
step: 220, loss: 0.006639664061367512
step: 230, loss: 0.007528066635131836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9820627802690582, f1=0.979591836734694,
best_f1=0.9842342342342343
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021785960416309536
step: 10, loss: 0.00040250312304124236
step: 20, loss: 0.12403469532728195
step: 30, loss: 0.00041793289710767567
step: 40, loss: 0.00039703049696981907
step: 50, loss: 0.0017814399907365441
step: 60, loss: 0.002653210423886776

```

```

step: 70, loss: 0.06621233373880386
step: 80, loss: 0.0016771507216617465
step: 90, loss: 0.06418151408433914
step: 100, loss: 0.00011606374027905986
step: 110, loss: 0.00015473426901735365
step: 120, loss: 0.007146098650991917
step: 130, loss: 0.0003944797208532691
step: 140, loss: 0.0013642878038808703
step: 150, loss: 0.00023450936714652926
step: 160, loss: 0.00047648866893723607
step: 170, loss: 0.00020325025252532214
step: 180, loss: 0.008449436165392399
step: 190, loss: 0.0053817881271243095
step: 200, loss: 0.0006297044456005096
step: 210, loss: 0.001967457588762045
step: 220, loss: 0.0001784659834811464
step: 230, loss: 0.00024879787815734744
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9864559819413092, f1=0.984090909090909,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016129047435242683
step: 10, loss: 0.00015630151028744876
step: 20, loss: 0.0037590127903968096
step: 30, loss: 8.345271635334939e-05
step: 40, loss: 0.0016175901982933283
step: 50, loss: 0.01635199412703514
step: 60, loss: 0.00016275691450573504
step: 70, loss: 0.0002519915287848562
step: 80, loss: 0.00011302579514449462
step: 90, loss: 7.993293547770008e-05
step: 100, loss: 0.00023692422837484628
step: 110, loss: 0.0023727002553641796
step: 120, loss: 0.00010575114720268175
step: 130, loss: 0.0001546186103951186
step: 140, loss: 0.00013102867524139583
step: 150, loss: 0.014341719448566437
step: 160, loss: 0.0001071694350684993
step: 170, loss: 0.0001387877855449915
step: 180, loss: 0.0036098239943385124

```

```

step: 190, loss: 0.0010047554969787598
step: 200, loss: 0.0009127795347012579
step: 210, loss: 0.15636050701141357
step: 220, loss: 0.0002305733214598149
step: 230, loss: 0.010273666121065617
/usr/local/lib/python3.7/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=0.9840546697038726, f1=0.9817767653758542,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007425214862450957
step: 10, loss: 0.0024188512470573187
step: 20, loss: 0.0005368830752559006
step: 30, loss: 0.0003530595568008721
step: 40, loss: 5.087699901196174e-05
step: 50, loss: 0.00020003062672913074
step: 60, loss: 0.0015509941149502993
step: 70, loss: 0.0018888169433921576
step: 80, loss: 0.0003071664832532406
step: 90, loss: 0.005168256349861622
step: 100, loss: 0.00038361220504157245
step: 110, loss: 0.006304699927568436
step: 120, loss: 0.00039827165892347693
step: 130, loss: 0.00011000464292010292
step: 140, loss: 0.000166057056048885
step: 150, loss: 0.0002025776047958061
step: 160, loss: 0.048595547676086426
step: 170, loss: 0.001967905554920435
step: 180, loss: 0.00015581384650431573
step: 190, loss: 0.00017461962124798447
step: 200, loss: 0.0051185558550059795
step: 210, loss: 9.550434333505109e-05
step: 220, loss: 0.0024597805459052324
step: 230, loss: 0.0024708088021725416
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9852440408626559,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003154588339384645
step: 10, loss: 0.0006393385701812804
step: 20, loss: 0.003098512766882777

```

```

step: 30, loss: 0.0334937646985054
step: 40, loss: 0.00015628428081981838
step: 50, loss: 0.00031886951182968915
step: 60, loss: 0.001052479026839137
step: 70, loss: 0.001589268445968628
step: 80, loss: 0.00018820175318978727
step: 90, loss: 0.0025462552439421415
step: 100, loss: 0.00010471537098055705
step: 110, loss: 6.949093949515373e-05
step: 120, loss: 0.00011953273497056216
step: 130, loss: 0.0037367776967585087
step: 140, loss: 0.00012555537978187203
step: 150, loss: 0.00019579687796067446
step: 160, loss: 0.0009187937248498201
step: 170, loss: 0.012379252351820469
step: 180, loss: 0.0002081599086523056
step: 190, loss: 0.0021726733539253473
step: 200, loss: 0.00396308908239007
step: 210, loss: 0.003049410181120038
step: 220, loss: 0.0005120654823258519
step: 230, loss: 0.0036688433028757572
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9796380090497738,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027725850231945515
step: 10, loss: 5.430261080618948e-05
step: 20, loss: 0.0006476076669059694
step: 30, loss: 4.797057408723049e-05
step: 40, loss: 7.12958790245466e-05
step: 50, loss: 0.0004190905310679227
step: 60, loss: 0.008875411935150623
step: 70, loss: 0.0003464212059043348
step: 80, loss: 0.00018523013568483293
step: 90, loss: 4.9835321988211945e-05
step: 100, loss: 0.0004642051353584975
step: 110, loss: 0.005275449715554714
step: 120, loss: 7.66482698963955e-05
step: 130, loss: 8.054973295656964e-05
step: 140, loss: 5.8280227676732466e-05
step: 150, loss: 7.892078428994864e-05
step: 160, loss: 0.0003503094194456935
step: 170, loss: 0.00040834309766069055
step: 180, loss: 0.01752229779958725

```

```

step: 190, loss: 0.0004982220707461238
step: 200, loss: 3.193755037500523e-05
step: 210, loss: 0.0006613027071580291
step: 220, loss: 0.00014356324390973896
step: 230, loss: 0.0006235099281184375
/usr/local/lib/python3.7/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=0.9852774631936579, f1=0.9829351535836178,
best_f1=0.984090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012661839718930423
step: 10, loss: 7.193889177870005e-05
step: 20, loss: 0.001347532612271607
step: 30, loss: 0.00036243474460206926
step: 40, loss: 0.000748537655454129
step: 50, loss: 6.290247983997688e-05
step: 60, loss: 8.110873750410974e-05
step: 70, loss: 0.15365830063819885
step: 80, loss: 0.00015803858696017414
step: 90, loss: 0.002506461227312684
step: 100, loss: 0.0003668049757834524
step: 110, loss: 0.00020050164312124252
step: 120, loss: 3.120899054920301e-05
step: 130, loss: 0.00028524361550807953
step: 140, loss: 7.155702769523486e-05
step: 150, loss: 9.854482777882367e-05
step: 160, loss: 0.002379852579906583
step: 170, loss: 0.00011521793931024149
step: 180, loss: 6.580066110473126e-05
step: 190, loss: 0.002628766931593418
step: 200, loss: 0.0002535637468099594
step: 210, loss: 0.00010808013757923618
step: 220, loss: 0.00012773652269970626
step: 230, loss: 5.8979905588785186e-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=0.9875706214689265, f1=0.9829738933030647,
best_f1=0.9829738933030647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004937507212162018
step: 10, loss: 0.0004324932233430445
step: 20, loss: 0.0037092517595738173
step: 30, loss: 8.902431727619842e-05
step: 40, loss: 4.2610117816366255e-05
step: 50, loss: 8.821905066724867e-05
step: 60, loss: 0.015965469181537628
step: 70, loss: 0.0002185619086958468
step: 80, loss: 5.986998075968586e-05
step: 90, loss: 9.399048576597124e-05
step: 100, loss: 3.59203404514119e-05
step: 110, loss: 0.004233133047819138
step: 120, loss: 0.0048450748436152935
step: 130, loss: 0.00345942797139287
step: 140, loss: 0.01232092548161745
step: 150, loss: 0.00018497955170460045
step: 160, loss: 0.0024385375436395407
step: 170, loss: 7.100641960278153e-05
step: 180, loss: 0.0003189168346580118
step: 190, loss: 0.00014379926142282784
step: 200, loss: 0.0016721318243071437
step: 210, loss: 0.01896514743566513
step: 220, loss: 0.00018268398707732558
step: 230, loss: 8.789252751739696e-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=0.9853107344632768, f1=0.9829738933030647,
best_f1=0.9829738933030647
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

keep_batchnorm_fp32      : True
master_weights           : True
loss_scale                : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:17, 143.85it/s]
load_f1 = 0.9864559819413092
real_f1 = 0.9864559819413092
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:32, 134.11it/s]

```

4.1.3 DBLP-GoogleScholar - Running the matcher

```

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

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

```

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
step: 0, loss: 0.617284893989563
```

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

```
step: 10, loss: 0.4782482385635376
```

```
step: 20, loss: 0.26416030526161194
```

```
step: 30, loss: 0.3762555718421936
```

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

```
step: 40, loss: 0.2739384174346924
```

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

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

```
step: 50, loss: 1.0934109687805176
```

```
step: 60, loss: 0.24329063296318054
```

step: 70, loss: 0.25844255089759827
step: 80, loss: 0.08348731696605682
step: 90, loss: 0.14782845973968506
step: 100, loss: 0.2858273983001709
step: 110, loss: 0.11756345629692078
step: 120, loss: 0.14633987843990326
step: 130, loss: 0.1415497362613678
step: 140, loss: 0.33209797739982605
step: 150, loss: 0.10859663039445877
step: 160, loss: 0.17688730359077454
step: 170, loss: 0.11697231233119965
step: 180, loss: 0.2103944569826126
step: 190, loss: 0.2411806732416153
step: 200, loss: 0.08068548887968063
step: 210, loss: 0.036091435700654984
step: 220, loss: 0.07114587724208832
step: 230, loss: 0.20119167864322662
step: 240, loss: 0.04284684360027313
step: 250, loss: 0.11961301416158676
step: 260, loss: 0.17455345392227173
step: 270, loss: 0.374035120010376
step: 280, loss: 0.13779973983764648
step: 290, loss: 0.1344231367111206
step: 300, loss: 0.161468505859375
step: 310, loss: 0.1975269466638565
step: 320, loss: 0.061528582125902176
step: 330, loss: 0.05920959636569023
step: 340, loss: 0.4707822799682617
step: 350, loss: 0.2059343159198761
step: 360, loss: 0.07145454734563828
step: 370, loss: 0.03629399091005325
step: 380, loss: 0.13232187926769257
step: 390, loss: 0.04798588901758194
step: 400, loss: 0.05557509511709213
step: 410, loss: 0.2910923957824707
step: 420, loss: 0.056954339146614075
step: 430, loss: 0.10943325608968735
step: 440, loss: 0.08835215121507645
step: 450, loss: 0.1162901371717453
step: 460, loss: 0.0774737223982811
step: 470, loss: 0.1139431893825531
step: 480, loss: 0.22768770158290863
step: 490, loss: 0.2651587724685669
step: 500, loss: 0.047581352293491364
step: 510, loss: 0.04038761928677559
step: 520, loss: 0.12636622786521912
step: 530, loss: 0.04483238607645035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.902405810258738, f1=0.9009090909090909,
best_f1=0.9009090909090909
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17268840968608856
step: 10, loss: 0.08897114545106888
step: 20, loss: 0.05880781635642052
step: 30, loss: 0.1566813439130783
step: 40, loss: 0.0342981182038784
step: 50, loss: 0.1337858885526657
step: 60, loss: 0.1265101432800293
step: 70, loss: 0.1586204469203949
step: 80, loss: 0.11411997675895691
step: 90, loss: 0.04341376945376396
step: 100, loss: 0.18650870025157928
step: 110, loss: 0.01991661824285984
step: 120, loss: 0.12764319777488708
step: 130, loss: 0.07333475351333618
step: 140, loss: 0.06699039041996002
step: 150, loss: 0.11399953067302704
step: 160, loss: 0.04118715599179268
step: 170, loss: 0.051527105271816254
step: 180, loss: 0.021843433380126953
step: 190, loss: 0.03973539173603058
step: 200, loss: 0.25730112195014954
step: 210, loss: 0.1102830246090889
step: 220, loss: 0.0038650077767670155
step: 230, loss: 0.10369934141635895
step: 240, loss: 0.04512151703238487
step: 250, loss: 0.03201952949166298
step: 260, loss: 0.1500484198331833
step: 270, loss: 0.03312751650810242
step: 280, loss: 0.060651738196611404
step: 290, loss: 0.1462966650724411
step: 300, loss: 0.059456534683704376
step: 310, loss: 0.11549657583236694
step: 320, loss: 0.1504022628068924
step: 330, loss: 0.07149215042591095
step: 340, loss: 0.30333590507507324
step: 350, loss: 0.08377712219953537
step: 360, loss: 0.12130790203809738

```

```

step: 370, loss: 0.007836582139134407
step: 380, loss: 0.12473524361848831
step: 390, loss: 0.021747518330812454
step: 400, loss: 0.08214353770017624
step: 410, loss: 0.10551035404205322
step: 420, loss: 0.12238747626543045
step: 430, loss: 0.21629106998443604
step: 440, loss: 0.02836407534778118
step: 450, loss: 0.06698336452245712
step: 460, loss: 0.12730246782302856
step: 470, loss: 0.138342022895813
step: 480, loss: 0.031239259988069534
step: 490, loss: 0.2398652881383896
step: 500, loss: 0.005841784179210663
step: 510, loss: 0.06795677542686462
step: 520, loss: 0.4165225028991699
step: 530, loss: 0.0658244714140892
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9192660550458716, f1=0.9187471629596006,
best_f1=0.9187471629596006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.150130033493042
step: 10, loss: 0.0314149335026741
step: 20, loss: 0.024425925686955452
step: 30, loss: 0.09067873656749725
step: 40, loss: 0.03768138214945793
step: 50, loss: 0.09324091672897339
step: 60, loss: 0.061903856694698334
step: 70, loss: 0.013112800195813179
step: 80, loss: 0.2578161656856537
step: 90, loss: 0.03255307674407959
step: 100, loss: 0.15077190101146698
step: 110, loss: 0.0723617747426033
step: 120, loss: 0.1358250081539154
step: 130, loss: 0.03988385200500488
step: 140, loss: 0.007366796489804983
step: 150, loss: 0.023759102448821068
step: 160, loss: 0.027168385684490204
step: 170, loss: 0.01638917811214924
step: 180, loss: 0.056700751185417175

```

```

step: 190, loss: 0.008512300439178944
step: 200, loss: 0.09656774252653122
step: 210, loss: 0.08446631580591202
step: 220, loss: 0.017910132184624672
step: 230, loss: 0.07040873169898987
step: 240, loss: 0.08308633416891098
step: 250, loss: 0.12170187383890152
step: 260, loss: 0.059775400906801224
step: 270, loss: 0.007328015752136707
step: 280, loss: 0.015342137776315212
step: 290, loss: 0.020687632262706757
step: 300, loss: 0.2688521146774292
step: 310, loss: 0.17679083347320557
step: 320, loss: 0.00703124376013875
step: 330, loss: 0.0026813684962689877
step: 340, loss: 0.019692936912178993
step: 350, loss: 0.1843792349100113
step: 360, loss: 0.028090955689549446
step: 370, loss: 0.07734725624322891
step: 380, loss: 0.11022278666496277
step: 390, loss: 0.05685867369174957
step: 400, loss: 0.18076737225055695
step: 410, loss: 0.03887759894132614
step: 420, loss: 0.10739246010780334
step: 430, loss: 0.04649639129638672
step: 440, loss: 0.23329459130764008
step: 450, loss: 0.17680050432682037
step: 460, loss: 0.059122879058122635
step: 470, loss: 0.017363250255584717
step: 480, loss: 0.200330451130867
step: 490, loss: 0.03886551037430763
step: 500, loss: 0.05955536291003227
step: 510, loss: 0.0581543892621994
step: 520, loss: 0.01138918288052082
step: 530, loss: 0.03109051287174225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9208566108007449, f1=0.912739150723285,
best_f1=0.912739150723285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05033033341169357

```

step: 10, loss: 0.009373889304697514
step: 20, loss: 0.12361646443605423
step: 30, loss: 0.02361442893743515
step: 40, loss: 0.022961929440498352
step: 50, loss: 0.03570215776562691
step: 60, loss: 0.0045125423930585384
step: 70, loss: 0.032733410596847534
step: 80, loss: 0.15621614456176758
step: 90, loss: 0.09164389222860336
step: 100, loss: 0.005166068207472563
step: 110, loss: 0.05998534709215164
step: 120, loss: 0.04375624656677246
step: 130, loss: 0.07668532431125641
step: 140, loss: 0.1305665224790573
step: 150, loss: 0.04442906752228737
step: 160, loss: 0.021223751828074455
step: 170, loss: 0.011218964122235775
step: 180, loss: 0.05121619626879692
step: 190, loss: 0.09152851998806
step: 200, loss: 0.05709437280893326
step: 210, loss: 0.011368729174137115
step: 220, loss: 0.025127733126282692
step: 230, loss: 0.02273300103843212
step: 240, loss: 0.007052634377032518
step: 250, loss: 0.05506129562854767
step: 260, loss: 0.0023515745997428894
step: 270, loss: 0.17676928639411926
step: 280, loss: 0.018789982423186302
step: 290, loss: 0.03983933478593826
step: 300, loss: 0.03731038048863411
step: 310, loss: 0.03782513365149498
step: 320, loss: 0.17739567160606384
step: 330, loss: 0.009925315156579018
step: 340, loss: 0.003854963928461075
step: 350, loss: 0.1962089091539383
step: 360, loss: 0.06561867892742157
step: 370, loss: 0.01496550627052784
step: 380, loss: 0.005120525136590004
step: 390, loss: 0.0036921128630638123
step: 400, loss: 0.005258080083876848
step: 410, loss: 0.024025658145546913
step: 420, loss: 0.015275309793651104
step: 430, loss: 0.011382227763533592
step: 440, loss: 0.011120524257421494
step: 450, loss: 0.12624305486679077
step: 460, loss: 0.0813184380531311
step: 470, loss: 0.006204581819474697
step: 480, loss: 0.004215720575302839


```

step: 490, loss: 0.01452107634395361
step: 500, loss: 0.11615893244743347
step: 510, loss: 0.07451581209897995
step: 520, loss: 0.013263002969324589
step: 530, loss: 0.045820172876119614
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9302112029384756, f1=0.9231473010064044,
best_f1=0.9231473010064044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006927515845745802
step: 10, loss: 0.01601782627403736
step: 20, loss: 0.009186084382236004
step: 30, loss: 0.1691954880952835
step: 40, loss: 0.011237859725952148
step: 50, loss: 0.07003667205572128
step: 60, loss: 0.021835096180438995
step: 70, loss: 0.004522776696830988
step: 80, loss: 0.00812008511275053
step: 90, loss: 0.025008831173181534
step: 100, loss: 0.049243975430727005
step: 110, loss: 0.006102921906858683
step: 120, loss: 0.1743173450231552
step: 130, loss: 0.0038408413529396057
step: 140, loss: 0.011179951019585133
step: 150, loss: 0.02607707493007183
step: 160, loss: 0.12838232517242432
step: 170, loss: 0.03452931344509125
step: 180, loss: 0.007559336721897125
step: 190, loss: 0.005091474391520023
step: 200, loss: 0.0034642298705875874
step: 210, loss: 0.10932209342718124
step: 220, loss: 0.019235901534557343
step: 230, loss: 0.020300429314374924
step: 240, loss: 0.008342843502759933
step: 250, loss: 0.07067295163869858
step: 260, loss: 0.014347872696816921
step: 270, loss: 0.0030306524131447077
step: 280, loss: 0.11619210243225098
step: 290, loss: 0.040917254984378815
step: 300, loss: 0.20227758586406708

```

```

step: 310, loss: 0.0837261751294136
step: 320, loss: 0.09018533676862717
step: 330, loss: 0.023619357496500015
step: 340, loss: 0.011138034984469414
step: 350, loss: 0.025123801082372665
step: 360, loss: 0.0008549454505555332
step: 370, loss: 0.0011741707567125559
step: 380, loss: 0.002104658167809248
step: 390, loss: 0.06895823776721954
step: 400, loss: 0.11991286277770996
step: 410, loss: 0.06075707823038101
step: 420, loss: 0.19273002445697784
step: 430, loss: 0.04136325418949127
step: 440, loss: 0.0040584770031273365
step: 450, loss: 0.0016299603739753366
step: 460, loss: 0.030864309519529343
step: 470, loss: 0.0513264462351799
step: 480, loss: 0.03631729632616043
step: 490, loss: 0.028187397867441177
step: 500, loss: 0.035610396414995193
step: 510, loss: 0.02128816582262516
step: 520, loss: 0.27299779653549194
step: 530, loss: 0.017031561583280563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.932963476652797, f1=0.9314179796107508,
best_f1=0.9314179796107508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0251386146992445
step: 10, loss: 0.0057729026302695274
step: 20, loss: 0.00846484862267971
step: 30, loss: 0.0013159915106371045
step: 40, loss: 0.05818937346339226
step: 50, loss: 0.026824288070201874
step: 60, loss: 0.03099253587424755
step: 70, loss: 0.004268394783139229
step: 80, loss: 0.03770576789975166
step: 90, loss: 0.12578411400318146
step: 100, loss: 0.05389925092458725
step: 110, loss: 0.021039795130491257
step: 120, loss: 0.04966992512345314

```

```

step: 130, loss: 0.0032676421105861664
step: 140, loss: 0.0336383692920208
step: 150, loss: 0.01015594881027937
step: 160, loss: 0.050899941474199295
step: 170, loss: 0.0024933863896876574
step: 180, loss: 0.008803987875580788
step: 190, loss: 0.16579844057559967
step: 200, loss: 0.1505102962255478
step: 210, loss: 0.004796930588781834
step: 220, loss: 0.027705969288945198
step: 230, loss: 0.07387477159500122
step: 240, loss: 0.10668005049228668
step: 250, loss: 0.05267135053873062
step: 260, loss: 0.021744804456830025
step: 270, loss: 0.010953249409794807
step: 280, loss: 0.1050516813993454
step: 290, loss: 0.07822467386722565
step: 300, loss: 0.004239434376358986
step: 310, loss: 0.059788789600133896
step: 320, loss: 0.002813818631693721
step: 330, loss: 0.01549531053751707
step: 340, loss: 0.001067449222318828
step: 350, loss: 0.0055549475364387035
step: 360, loss: 0.026470327749848366
step: 370, loss: 0.013340862467885017
step: 380, loss: 0.001798618002794683
step: 390, loss: 0.003585513448342681
step: 400, loss: 0.05383109673857689
step: 410, loss: 0.002912794006988406
step: 420, loss: 0.004142374731600285
step: 430, loss: 0.00045239069731906056
step: 440, loss: 0.0016848387895151973
step: 450, loss: 0.175912544131279
step: 460, loss: 0.011207680217921734
step: 470, loss: 0.0017661906313151121
step: 480, loss: 0.012238946743309498
step: 490, loss: 0.007146426010876894
step: 500, loss: 0.001662468770518899
step: 510, loss: 0.26738229393959045
step: 520, loss: 0.0003026924387086183
step: 530, loss: 0.0021439732518047094
/usr/local/lib/python3.7/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=0.9286043298019346, f1=0.9174825174825175,
best_f1=0.9314179796107508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003126990981400013
step: 10, loss: 0.025847947224974632
step: 20, loss: 0.008445674553513527
step: 30, loss: 0.02131718397140503
step: 40, loss: 0.007984454743564129
step: 50, loss: 0.015561887994408607
step: 60, loss: 0.002972874790430069
step: 70, loss: 0.0005530305206775665
step: 80, loss: 0.07013770192861557
step: 90, loss: 0.0008748235995881259
step: 100, loss: 0.004403166007250547
step: 110, loss: 0.0032700528390705585
step: 120, loss: 0.007362982723861933
step: 130, loss: 0.014032007195055485
step: 140, loss: 0.011053254827857018
step: 150, loss: 0.011300662532448769
step: 160, loss: 0.0028897763695567846
step: 170, loss: 0.010036535561084747
step: 180, loss: 0.10986049473285675
step: 190, loss: 0.0479092039167881
step: 200, loss: 0.002112818881869316
step: 210, loss: 0.001654896535910666
step: 220, loss: 0.046246081590652466
step: 230, loss: 0.0009951061801984906
step: 240, loss: 0.0015333350747823715
step: 250, loss: 0.020488496869802475
step: 260, loss: 0.0022479002363979816
step: 270, loss: 0.01636660285294056
step: 280, loss: 0.0055391788482666016
step: 290, loss: 0.01459481380879879
step: 300, loss: 0.003379348898306489
step: 310, loss: 0.07356192171573639
step: 320, loss: 0.007871480658650398
step: 330, loss: 0.0010709615889936686
step: 340, loss: 0.0016459182370454073
step: 350, loss: 0.00863460823893547
step: 360, loss: 0.00558296637609601
step: 370, loss: 0.004046321380883455
step: 380, loss: 0.0011179291177541018
step: 390, loss: 0.02620280534029007
step: 400, loss: 0.015949856489896774
step: 410, loss: 0.008249735459685326
step: 420, loss: 0.006872677709907293
step: 430, loss: 0.04505356401205063
step: 440, loss: 0.010788715444505215
step: 450, loss: 0.26719415187835693
step: 460, loss: 0.005323857069015503
```

```

step: 470, loss: 0.1582753211259842
step: 480, loss: 0.01939881592988968
step: 490, loss: 0.00498679606243968
step: 500, loss: 0.010003272444009781
step: 510, loss: 0.0005342514486983418
step: 520, loss: 0.005458083003759384
step: 530, loss: 0.002290351316332817
/usr/local/lib/python3.7/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=0.9215955983493811, f1=0.9092592592592593,
best_f1=0.9314179796107508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000960883277002722
step: 10, loss: 0.008751815184950829
step: 20, loss: 0.0008988957852125168
step: 30, loss: 0.006863206159323454
step: 40, loss: 0.006608037743717432
step: 50, loss: 0.05232347548007965
step: 60, loss: 0.001079275505617261
step: 70, loss: 0.007398138288408518
step: 80, loss: 0.005146307405084372
step: 90, loss: 0.027069615200161934
step: 100, loss: 0.011090937070548534
step: 110, loss: 0.0022715460509061813
step: 120, loss: 0.0039028034079819918
step: 130, loss: 0.0073945242911577225
step: 140, loss: 0.002104600891470909
step: 150, loss: 0.004275529645383358
step: 160, loss: 0.014258806593716145
step: 170, loss: 0.03688448667526245
step: 180, loss: 0.005894681438803673
step: 190, loss: 0.1110907793045044
step: 200, loss: 0.012748047709465027
step: 210, loss: 0.05194050073623657
step: 220, loss: 0.015756191685795784
step: 230, loss: 0.13396646082401276
step: 240, loss: 0.03495645895600319
step: 250, loss: 0.0013907711254432797
step: 260, loss: 0.005435395520180464
step: 270, loss: 0.05292563512921333
step: 280, loss: 0.0009389918413944542
step: 290, loss: 0.03542700409889221
step: 300, loss: 0.0024252873845398426
step: 310, loss: 0.0014204145409166813
step: 320, loss: 0.0038502365350723267

```

```

step: 330, loss: 0.013034873642027378
step: 340, loss: 0.010962097905576229
step: 350, loss: 0.0028318173717707396
step: 360, loss: 0.030214639380574226
step: 370, loss: 0.16304516792297363
step: 380, loss: 0.0004678385448642075
step: 390, loss: 0.0025676912628114223
step: 400, loss: 0.03234437480568886
step: 410, loss: 0.006061968393623829
step: 420, loss: 0.0005152482772246003
step: 430, loss: 0.00317495409399271
step: 440, loss: 0.08541222661733627
step: 450, loss: 0.0030380531679838896
step: 460, loss: 0.036524392664432526
step: 470, loss: 0.031705062836408615
step: 480, loss: 0.03659329563379288
step: 490, loss: 0.006825421005487442
step: 500, loss: 0.0033034367952495813
step: 510, loss: 0.07784533500671387
step: 520, loss: 0.0007455720915459096
step: 530, loss: 0.016583019867539406
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9269662921348314, f1=0.9199063231850116,
best_f1=0.9314179796107508
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005185039481148124
step: 10, loss: 0.00573626859113574
step: 20, loss: 0.0012510053347796202
step: 30, loss: 0.19032733142375946
step: 40, loss: 0.0023616435937583447
step: 50, loss: 0.001021567964926362
step: 60, loss: 0.00032683450262993574
step: 70, loss: 0.0022714051883667707
step: 80, loss: 0.0038024389650672674
step: 90, loss: 0.0374935120344162
step: 100, loss: 0.0013368877116590738
step: 110, loss: 0.018323004245758057
step: 120, loss: 0.0029490827582776546
step: 130, loss: 0.0006655059405602515
step: 140, loss: 0.0006133682327345014
step: 150, loss: 0.0031699086539447308
step: 160, loss: 0.001524075516499579
step: 170, loss: 0.009833212941884995

```

```

step: 180, loss: 0.0016673324862495065
step: 190, loss: 0.0005782920052297413
step: 200, loss: 0.05501784756779671
step: 210, loss: 0.14902816712856293
step: 220, loss: 0.004255413077771664
step: 230, loss: 0.0004896405152976513
step: 240, loss: 0.0017327357782050967
step: 250, loss: 0.042819101363420486
step: 260, loss: 0.03536440059542656
step: 270, loss: 0.00944449845701456
step: 280, loss: 0.01343455445021391
step: 290, loss: 0.0014820712385699153
step: 300, loss: 0.002620021114125848
step: 310, loss: 0.0078084515407681465
step: 320, loss: 0.005134719889611006
step: 330, loss: 0.012331499718129635
step: 340, loss: 0.03410129249095917
step: 350, loss: 0.06584448367357254
step: 360, loss: 0.0011448776349425316
step: 370, loss: 0.004562468733638525
step: 380, loss: 0.0007956580957397819
step: 390, loss: 6.141741323517635e-05
step: 400, loss: 0.0583292692899704
step: 410, loss: 0.001276268158107996
step: 420, loss: 0.002065780572593212
step: 430, loss: 0.004231240600347519
step: 440, loss: 0.00019630538008641452
step: 450, loss: 0.005229976959526539
step: 460, loss: 0.00043396951514296234
step: 470, loss: 0.0006998951430432498
step: 480, loss: 0.0007751507801003754
step: 490, loss: 0.008083074353635311
step: 500, loss: 0.002214076928794384
step: 510, loss: 0.008102747611701488
step: 520, loss: 0.005369021091610193
step: 530, loss: 0.11553199589252472
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9338919925512105, f1=0.9255121042830541,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.005992721766233444
step: 10, loss: 0.002959729405120015
step: 20, loss: 0.0008037597872316837
step: 30, loss: 0.021741045638918877
step: 40, loss: 0.00018528154760133475
step: 50, loss: 0.0002935800876002759
step: 60, loss: 0.004720487166196108
step: 70, loss: 0.007121972739696503
step: 80, loss: 0.0021480892319232225
step: 90, loss: 0.008022544905543327
step: 100, loss: 0.0025052570272237062
step: 110, loss: 0.006111218128353357
step: 120, loss: 0.00027984281769022346
step: 130, loss: 0.08065585047006607
step: 140, loss: 0.0030572949908673763
step: 150, loss: 0.071358822286129
step: 160, loss: 0.01598622463643551
step: 170, loss: 0.003937490750104189
step: 180, loss: 0.019649315625429153
step: 190, loss: 0.002303289482370019
step: 200, loss: 0.003943735733628273
step: 210, loss: 0.015050258487462997
step: 220, loss: 0.0011115869274362922
step: 230, loss: 0.0009410008206032217
step: 240, loss: 0.0016423598863184452
step: 250, loss: 0.01815556176006794
step: 260, loss: 0.004198779352009296
step: 270, loss: 0.000373553455574438
step: 280, loss: 0.008798115886747837
step: 290, loss: 0.009173594415187836
step: 300, loss: 0.05685081705451012
step: 310, loss: 0.014910036697983742
step: 320, loss: 0.05782052502036095
step: 330, loss: 0.0036791954189538956
step: 340, loss: 0.004030929878354073
step: 350, loss: 0.00041930811130441725
step: 360, loss: 0.00036597560392692685
step: 370, loss: 0.0001606543082743883
step: 380, loss: 0.001847034553065896
step: 390, loss: 0.0005187420174479485
step: 400, loss: 0.001469831564463675
step: 410, loss: 0.0003560897312127054
step: 420, loss: 0.0007333728135563433
step: 430, loss: 0.0008061093394644558
step: 440, loss: 0.00025405114865861833
step: 450, loss: 0.2055104523897171
step: 460, loss: 0.0005068997270427644
step: 470, loss: 0.001563475583679974


```

step: 480, loss: 0.0022912088315933943
step: 490, loss: 0.0011648096842691302
step: 500, loss: 0.0010877561289817095
step: 510, loss: 0.011052446439862251
step: 520, loss: 0.0003902667958755046
step: 530, loss: 0.01265582162886858
/usr/local/lib/python3.7/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=0.9305164319248826, f1=0.928772258669166,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013146920537110418
step: 10, loss: 0.0005849735462106764
step: 20, loss: 0.0013305684551596642
step: 30, loss: 0.0003890228399541229
step: 40, loss: 0.0383707620203495
step: 50, loss: 0.00044115117634646595
step: 60, loss: 0.008306773379445076
step: 70, loss: 0.00036437282687984407
step: 80, loss: 0.0006653191521763802
step: 90, loss: 0.017144758254289627
step: 100, loss: 0.00026783309294842184
step: 110, loss: 0.00025730181368999183
step: 120, loss: 0.013373693451285362
step: 130, loss: 0.0004664890584535897
step: 140, loss: 0.001115297549404204
step: 150, loss: 0.0012438817648217082
step: 160, loss: 0.0014078021049499512
step: 170, loss: 0.09832889586687088
step: 180, loss: 0.0007477639592252672
step: 190, loss: 0.00022522482322528958
step: 200, loss: 0.0007285320316441357
step: 210, loss: 0.000796284235548228
step: 220, loss: 0.015829939395189285
step: 230, loss: 0.0002301548229297623
step: 240, loss: 0.003089808626100421
step: 250, loss: 0.00017669916269369423
step: 260, loss: 0.002516944659873843
step: 270, loss: 0.003756581572815776
step: 280, loss: 0.00026206308393739164
step: 290, loss: 0.00017730161198414862
step: 300, loss: 0.00015649183478672057
step: 310, loss: 0.00037948210956528783
step: 320, loss: 0.0011479745153337717
step: 330, loss: 2.200473682023585e-05

```

```

step: 340, loss: 0.002350603463128209
step: 350, loss: 0.00013304472668096423
step: 360, loss: 0.003943199757486582
step: 370, loss: 0.002156467642635107
step: 380, loss: 0.01808573305606842
step: 390, loss: 0.0011705500073730946
step: 400, loss: 0.0004869074618909508
step: 410, loss: 0.0008626448106952012
step: 420, loss: 0.0009019009885378182
step: 430, loss: 0.0020761052146553993
step: 440, loss: 9.931746171787381e-05
step: 450, loss: 0.0009267627028748393
step: 460, loss: 0.05432705581188202
step: 470, loss: 0.0020600208081305027
step: 480, loss: 0.0006096953293308616
step: 490, loss: 0.00017903988191392273
step: 500, loss: 0.0002570881333667785
step: 510, loss: 0.0015519977314397693
step: 520, loss: 0.0005849187145940959
step: 530, loss: 0.004614701960235834
/usr/local/lib/python3.7/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=0.9244402985074627, f1=0.9214417744916821,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006263298913836479
step: 10, loss: 0.000714750902261585
step: 20, loss: 0.0017466646386310458
step: 30, loss: 0.03130101040005684
step: 40, loss: 0.005644228775054216
step: 50, loss: 0.00016243643767666072
step: 60, loss: 0.00011756987078115344
step: 70, loss: 0.003985977731645107
step: 80, loss: 0.10755594074726105
step: 90, loss: 0.003949123900383711
step: 100, loss: 0.007019927725195885
step: 110, loss: 0.000842335051856935
step: 120, loss: 0.0018181328196078539
step: 130, loss: 0.0009278818033635616
step: 140, loss: 0.0003269758017268032
step: 150, loss: 0.007650346960872412
step: 160, loss: 0.011082588694989681
step: 170, loss: 0.0005804415559396148
step: 180, loss: 0.0003351195773575455
step: 190, loss: 0.001573605346493423

```

```

step: 200, loss: 0.00266434857621789
step: 210, loss: 0.0004230165504850447
step: 220, loss: 0.0001759290898917243
step: 230, loss: 0.0009528547525405884
step: 240, loss: 0.0002785473479889333
step: 250, loss: 0.0006033890531398356
step: 260, loss: 0.005990173667669296
step: 270, loss: 0.027417218312621117
step: 280, loss: 0.002966279862448573
step: 290, loss: 0.0006324401474557817
step: 300, loss: 9.542056068312377e-05
step: 310, loss: 0.0001662579452386126
step: 320, loss: 0.20808792114257812
step: 330, loss: 0.017903784289956093
step: 340, loss: 7.044738595141098e-05
step: 350, loss: 0.0003370067570358515
step: 360, loss: 0.0006736614159308374
step: 370, loss: 0.0009513969998806715
step: 380, loss: 5.6059590860968456e-05
step: 390, loss: 0.0069389729760587215
step: 400, loss: 0.0004068332491442561
step: 410, loss: 0.0007890319684520364
step: 420, loss: 0.0004046939720865339
step: 430, loss: 0.0005968661862425506
step: 440, loss: 0.004741601645946503
step: 450, loss: 0.002436507260426879
step: 460, loss: 9.632237197365612e-05
step: 470, loss: 0.005259502213448286
step: 480, loss: 0.0012895534746348858
step: 490, loss: 0.000559850363060832
step: 500, loss: 0.0001546875573694706
step: 510, loss: 0.004658810328692198
step: 520, loss: 0.0015437341062352061
step: 530, loss: 0.0011652217945083976
/usr/local/lib/python3.7/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=0.9330855018587362, f1=0.9302540415704387,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013044763181824237
step: 10, loss: 0.000263387308223173
step: 20, loss: 0.0007098221103660762
step: 30, loss: 7.786018977640197e-05
step: 40, loss: 0.02875479683279991
step: 50, loss: 0.0035058134235441685

```

step: 60, loss: 0.0009258160716854036
step: 70, loss: 0.004982514772564173
step: 80, loss: 0.009490151889622211
step: 90, loss: 0.0007718025590293109
step: 100, loss: 0.00011377550254110247
step: 110, loss: 0.0007242752471938729
step: 120, loss: 0.0001996014325413853
step: 130, loss: 0.0004580815730150789
step: 140, loss: 0.0004967297427356243
step: 150, loss: 0.003962190821766853
step: 160, loss: 0.007874323055148125
step: 170, loss: 0.009197228588163853
step: 180, loss: 0.00018789147725328803
step: 190, loss: 0.0010363635374233127
step: 200, loss: 0.0003092560509685427
step: 210, loss: 2.794132160488516e-05
step: 220, loss: 0.00012240390060469508
step: 230, loss: 0.008360073901712894
step: 240, loss: 0.0004099091747775674
step: 250, loss: 0.0010469924891367555
step: 260, loss: 0.0029625282622873783
step: 270, loss: 0.00023811090795788914
step: 280, loss: 0.0001053000014508143
step: 290, loss: 0.002978615928441286
step: 300, loss: 0.0036610509268939495
step: 310, loss: 5.087440513307229e-05
step: 320, loss: 2.073066934826784e-05
step: 330, loss: 0.0035896864719688892
step: 340, loss: 0.0016350928926840425
step: 350, loss: 0.00012605894880834967
step: 360, loss: 0.12892784178256989
step: 370, loss: 0.0010891204001381993
step: 380, loss: 0.00011460579844424501
step: 390, loss: 0.0011690425453707576
step: 400, loss: 0.001495064003393054
step: 410, loss: 7.505205576308072e-05
step: 420, loss: 0.0012957652797922492
step: 430, loss: 0.00026067078579217196
step: 440, loss: 0.0019298931583762169
step: 450, loss: 0.005753530189394951
step: 460, loss: 0.0013578905491158366
step: 470, loss: 0.10402494668960571
step: 480, loss: 9.573882380209398e-06
step: 490, loss: 9.879510616883636e-05
step: 500, loss: 0.008854140527546406
step: 510, loss: 0.00011923150304937735
step: 520, loss: 0.00018780262325890362
step: 530, loss: 4.9900369049282745e-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=0.9338235294117647, f1=0.9274303970789594,
best_f1=0.9255121042830541
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004216416738927364
step: 10, loss: 0.001035897177644074
step: 20, loss: 2.7520265575731173e-05
step: 30, loss: 0.0015041290316730738
step: 40, loss: 0.007422266528010368
step: 50, loss: 0.0018444570014253259
step: 60, loss: 0.0016771501395851374
step: 70, loss: 9.148837852990255e-05
step: 80, loss: 1.5422190699609928e-05
step: 90, loss: 3.461895539658144e-05
step: 100, loss: 2.3669243091717362e-05
step: 110, loss: 0.025403831154108047
step: 120, loss: 7.297797310457099e-06
step: 130, loss: 6.752800982212648e-05
step: 140, loss: 0.019921354949474335
step: 150, loss: 0.033069055527448654
step: 160, loss: 0.07184824347496033
step: 170, loss: 0.00366063485853374
step: 180, loss: 0.0006105338106863201
step: 190, loss: 0.0009912744862958789
step: 200, loss: 0.017923081293702126
step: 210, loss: 6.428138294722885e-05
step: 220, loss: 3.551621557562612e-05
step: 230, loss: 0.00720404414460063
step: 240, loss: 0.0008242945768870413
step: 250, loss: 0.003921110183000565
step: 260, loss: 4.9226495320908725e-05
step: 270, loss: 0.0009635213646106422
step: 280, loss: 0.014685237780213356
step: 290, loss: 0.001512891030870378
step: 300, loss: 0.0001232173526659608
step: 310, loss: 6.082585969124921e-05
step: 320, loss: 0.0001508271525381133
step: 330, loss: 0.004759150557219982
step: 340, loss: 0.0002828200231306255
step: 350, loss: 0.00017654159455560148
step: 360, loss: 0.0008864759001880884
step: 370, loss: 0.0019514778396114707
step: 380, loss: 0.007125849835574627
step: 390, loss: 9.921505261445418e-05
```

```

step: 400, loss: 0.009186021983623505
step: 410, loss: 0.0014865214470773935
step: 420, loss: 0.00013522320659831166
step: 430, loss: 0.008419918827712536
step: 440, loss: 0.0013576647033914924
step: 450, loss: 4.867386815021746e-05
step: 460, loss: 0.026691965758800507
step: 470, loss: 2.2040911062504165e-05
step: 480, loss: 0.001271312590688467
step: 490, loss: 0.004131829831749201
step: 500, loss: 0.0011976688401773572
step: 510, loss: 0.06857439875602722
step: 520, loss: 3.864612881443463e-05
step: 530, loss: 0.004223509691655636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9351503759398496, f1=0.9210649229332087,
best_f1=0.9210649229332087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.835950544977095e-05
step: 10, loss: 0.007853243499994278
step: 20, loss: 0.0016883236821740866
step: 30, loss: 0.012468945235013962
step: 40, loss: 0.00017634665709920228
step: 50, loss: 0.0010033881990239024
step: 60, loss: 0.009309557266533375
step: 70, loss: 9.256812336388975e-05
step: 80, loss: 0.00028507376555353403
step: 90, loss: 6.703443796141073e-05
step: 100, loss: 9.548242087475955e-05
step: 110, loss: 0.022612441331148148
step: 120, loss: 0.00011822906526504084
step: 130, loss: 0.0068045309744775295
step: 140, loss: 0.00013292233052197844
step: 150, loss: 8.409210568061098e-05
step: 160, loss: 0.0002237637381767854
step: 170, loss: 3.770170587813482e-05
step: 180, loss: 0.09714673459529877
step: 190, loss: 0.11786780506372452
step: 200, loss: 0.0012930898228660226
step: 210, loss: 0.003529164008796215

```

```

step: 220, loss: 0.045894477516412735
step: 230, loss: 0.002726073609665036
step: 240, loss: 0.028229400515556335
step: 250, loss: 0.0004715666873380542
step: 260, loss: 8.819693175610155e-05
step: 270, loss: 0.0011590052163228393
step: 280, loss: 0.0002864012203644961
step: 290, loss: 0.0006661395309492946
step: 300, loss: 0.0003182908985763788
step: 310, loss: 5.234772834228352e-05
step: 320, loss: 7.097314664861187e-05
step: 330, loss: 2.6075807909364812e-05
step: 340, loss: 6.431219662772492e-05
step: 350, loss: 0.002493486041203141
step: 360, loss: 0.0014670920791104436
step: 370, loss: 0.0004663038707803935
step: 380, loss: 0.0017647844506427646
step: 390, loss: 0.009272579103708267
step: 400, loss: 0.004301877226680517
step: 410, loss: 0.0011085941223427653
step: 420, loss: 0.0007119603687897325
step: 430, loss: 5.652654363075271e-05
step: 440, loss: 0.0027200602926313877
step: 450, loss: 0.0005932695930823684
step: 460, loss: 0.0010003396309912205
step: 470, loss: 1.1775492566812318e-05
step: 480, loss: 0.0017830788856372237
step: 490, loss: 0.00029169427580200136
step: 500, loss: 0.014552651904523373
step: 510, loss: 0.00010491106513654813
step: 520, loss: 9.433423838345334e-05
step: 530, loss: 5.541700011235662e-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=0.9353049907578558, f1=0.9225092250922509,
best_f1=0.9225092250922509
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 [00:23, 247.90it/s]
load_f1 = 0.9364107395195479
real_f1 = 0.9347111319868483
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:20, 211.82it/s]

```

4.1.4 Amazon-Google - Running the matcher

```

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

```



```

--save_model \
--dk general \
--dk product \
--summarize \
--da entry_swap \
--da attr_del

```

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

```
!mv *_dev.pt checkpoints/Amazon-Google.pt
```

Step 3: run the matcher

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

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

Downloading: 100% 899k/899k [00:00<00:00, 6.16MB/s]

Downloading: 100% 456k/456k [00:00<00:00, 3.66MB/s]

Downloading: 100% 501M/501M [00:08<00:00, 59.2MB/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: 0.4941563010215759
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.451974481344223
step: 20, loss: 0.4838372468948364
step: 30, loss: 0.307957261800766
step: 40, loss: 0.3020015358924866
step: 50, loss: 0.43260034918785095
step: 60, loss: 0.46466225385665894
step: 70, loss: 0.2989182770252228
step: 80, loss: 0.397663950920105
step: 90, loss: 0.2640441656112671
step: 100, loss: 0.24067182838916779
step: 110, loss: 0.2380286157131195
step: 120, loss: 0.34966355562210083
step: 130, loss: 0.24447320401668549
step: 140, loss: 0.5001601576805115
step: 150, loss: 0.40885958075523376
step: 160, loss: 0.5027796626091003
step: 170, loss: 0.2361486256122589
step: 180, loss: 0.3756333589553833
step: 190, loss: 0.6635289788246155
step: 200, loss: 0.3706989288330078
step: 210, loss: 0.4994335472583771
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3776681423187256
step: 10, loss: 0.18616126477718353
step: 20, loss: 0.49745863676071167
step: 30, loss: 0.4880135655403137
step: 40, loss: 0.4503121078014374
step: 50, loss: 0.24758809804916382
step: 60, loss: 0.30989038944244385
step: 70, loss: 0.4484732449054718
step: 80, loss: 0.31104475259780884
step: 90, loss: 0.3773651421070099
step: 100, loss: 0.5054131746292114
step: 110, loss: 0.3824058473110199
step: 120, loss: 0.24814960360527039
step: 130, loss: 0.16695843636989594
step: 140, loss: 0.23963862657546997
step: 150, loss: 0.4351697266101837
step: 160, loss: 0.16824764013290405
step: 170, loss: 0.61766517162323
step: 180, loss: 0.3156891465187073
step: 190, loss: 0.316119909286499
step: 200, loss: 0.14849363267421722
step: 210, loss: 0.31330224871635437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.18519984170953702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.25440433621406555
step: 10, loss: 0.23188944160938263
step: 20, loss: 0.46203377842903137
step: 30, loss: 0.2597634792327881
step: 40, loss: 0.44708895683288574
step: 50, loss: 0.46897977590560913
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 60, loss: 0.5521711111068726
step: 70, loss: 0.25863397121429443
step: 80, loss: 0.45650312304496765
step: 90, loss: 0.2399752140045166
step: 100, loss: 0.39342033863067627
step: 110, loss: 0.23340897262096405
step: 120, loss: 0.22276754677295685
step: 130, loss: 0.20112738013267517
step: 140, loss: 0.3533792197704315
step: 150, loss: 0.279423326253891

```

```

step: 160, loss: 0.1985575407743454
step: 170, loss: 0.3491940498352051
step: 180, loss: 0.23963838815689087
step: 190, loss: 0.1479991376399994
step: 200, loss: 0.25110796093940735
step: 210, loss: 0.30833131074905396
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2176015473887814, f1=0.2161119150988905,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.34919992089271545
step: 10, loss: 0.28826820850372314
step: 20, loss: 0.28257036209106445
step: 30, loss: 0.24041350185871124
step: 40, loss: 0.2232859581708908
step: 50, loss: 0.24460461735725403
step: 60, loss: 0.4591006338596344
step: 70, loss: 0.23471687734127045
step: 80, loss: 0.2263489067554474
step: 90, loss: 0.396145761013031
step: 100, loss: 0.4232441186904907
step: 110, loss: 0.6307476758956909
step: 120, loss: 0.3760738968849182
step: 130, loss: 0.6752288937568665
step: 140, loss: 0.5062586069107056
step: 150, loss: 0.38475292921066284
step: 160, loss: 0.31255224347114563
step: 170, loss: 0.14781507849693298
step: 180, loss: 0.11953137814998627
step: 190, loss: 0.16442395746707916
step: 200, loss: 0.30127450823783875
step: 210, loss: 0.38388678431510925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 4: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.3895803391933441
step: 10, loss: 0.3197759985923767
step: 20, loss: 0.30750900506973267
step: 30, loss: 0.24021480977535248
step: 40, loss: 0.44655466079711914
step: 50, loss: 0.3823348581790924
step: 60, loss: 0.3796880543231964
step: 70, loss: 0.24427439272403717
step: 80, loss: 0.4358472228050232
step: 90, loss: 0.4440169334411621
step: 100, loss: 0.23788577318191528
step: 110, loss: 0.16965503990650177
step: 120, loss: 0.23927856981754303
step: 130, loss: 0.3030185401439667
step: 140, loss: 0.5291065573692322
step: 150, loss: 0.3009544014930725
step: 160, loss: 0.23861615359783173
step: 170, loss: 0.3930393159389496
step: 180, loss: 0.23905400931835175
step: 190, loss: 0.5294354557991028
step: 200, loss: 0.3781454861164093
step: 210, loss: 0.23899757862091064
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15963837504386902
step: 10, loss: 0.3867270052433014
step: 20, loss: 0.3069915175437927
step: 30, loss: 0.233957901597023
step: 40, loss: 0.3134041428565979
step: 50, loss: 0.5989214181900024
step: 60, loss: 0.3196878433227539
step: 70, loss: 0.44483107328414917
step: 80, loss: 0.31694501638412476
step: 90, loss: 0.3140357732772827
step: 100, loss: 0.31792306900024414
step: 110, loss: 0.38168978691101074
step: 120, loss: 0.2395024299621582
step: 130, loss: 0.23471729457378387
step: 140, loss: 0.37480807304382324
step: 150, loss: 0.19326868653297424
step: 160, loss: 0.16108565032482147
step: 170, loss: 0.45016011595726013

```

```

step: 180, loss: 0.38357529044151306
step: 190, loss: 0.4432339668273926
step: 200, loss: 0.3249041736125946
step: 210, loss: 0.38162508606910706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4525177776813507
step: 10, loss: 0.3120381236076355
step: 20, loss: 0.4303344488143921
step: 30, loss: 0.2433795928955078
step: 40, loss: 0.25238773226737976
step: 50, loss: 0.37471553683280945
step: 60, loss: 0.3208082318305969
step: 70, loss: 0.24910517036914825
step: 80, loss: 0.5127520561218262
step: 90, loss: 0.37351176142692566
step: 100, loss: 0.4592723250389099
step: 110, loss: 0.3177390694618225
step: 120, loss: 0.3083241581916809
step: 130, loss: 0.4622890055179596
step: 140, loss: 0.30236610770225525
step: 150, loss: 0.30543145537376404
step: 160, loss: 0.5773932337760925
step: 170, loss: 0.6402401924133301
step: 180, loss: 0.2451862245798111
step: 190, loss: 0.24013663828372955
step: 200, loss: 0.3109624683856964
step: 210, loss: 0.30819758772850037
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6240262389183044
step: 10, loss: 0.30886131525039673
step: 20, loss: 0.2591104805469513
step: 30, loss: 0.2526475787162781
step: 40, loss: 0.23463204503059387
step: 50, loss: 0.16725057363510132

```

```

step: 60, loss: 0.1721634417772293
step: 70, loss: 0.4635390043258667
step: 80, loss: 0.31321093440055847
step: 90, loss: 0.377731591463089
step: 100, loss: 0.6266162991523743
step: 110, loss: 0.37847453355789185
step: 120, loss: 0.3227516710758209
step: 130, loss: 0.17266057431697845
step: 140, loss: 0.3190914988517761
step: 150, loss: 0.4575999975204468
step: 160, loss: 0.45663508772850037
step: 170, loss: 0.540723443031311
step: 180, loss: 0.2997058033943176
step: 190, loss: 0.24100443720817566
step: 200, loss: 0.4484339654445648
step: 210, loss: 0.3797999322414398
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5815706849098206
step: 10, loss: 0.2836182415485382
step: 20, loss: 0.37373456358909607
step: 30, loss: 0.16378848254680634
step: 40, loss: 0.1616528481245041
step: 50, loss: 0.4564427137374878
step: 60, loss: 0.17338347434997559
step: 70, loss: 0.3850531280040741
step: 80, loss: 0.3119693100452423
step: 90, loss: 0.3106548488140106
step: 100, loss: 0.47043758630752563
step: 110, loss: 0.5249307155609131
step: 120, loss: 0.38118839263916016
step: 130, loss: 0.24233251810073853
step: 140, loss: 0.6070509552955627
step: 150, loss: 0.11295387148857117
step: 160, loss: 0.38830313086509705
step: 170, loss: 0.31512340903282166
step: 180, loss: 0.4510428011417389
step: 190, loss: 0.30774733424186707
step: 200, loss: 0.30570754408836365
step: 210, loss: 0.30700892210006714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.18519984170953702, f1=0.18519984170953702,
best_f1=0.2161119150988905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3094070553779602
step: 10, loss: 0.3729322850704193
step: 20, loss: 0.25884535908699036
step: 30, loss: 0.16396932303905487
step: 40, loss: 0.3144213557243347
step: 50, loss: 0.4760274887084961
step: 60, loss: 0.20847897231578827
step: 70, loss: 0.3835851550102234
step: 80, loss: 0.12841413915157318
step: 90, loss: 0.36454877257347107
step: 100, loss: 0.36441633105278015
step: 110, loss: 0.12418817728757858
step: 120, loss: 0.5734175443649292
step: 130, loss: 0.21808677911758423
step: 140, loss: 0.32436272501945496
step: 150, loss: 0.24977882206439972
step: 160, loss: 0.27982521057128906
step: 170, loss: 0.15789487957954407
step: 180, loss: 0.3024680018424988
step: 190, loss: 0.19228404760360718
step: 200, loss: 0.4431131184101105
step: 210, loss: 0.22375407814979553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.38029782359679265, f1=0.3976470588235294,
best_f1=0.3976470588235294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24731293320655823
step: 10, loss: 0.2582995295524597
step: 20, loss: 0.2784358859062195
step: 30, loss: 0.18239964544773102
step: 40, loss: 0.3413184583187103
step: 50, loss: 0.407697856426239
step: 60, loss: 0.2544907331466675
step: 70, loss: 0.20386433601379395

```



```

step: 80, loss: 0.3586225211620331
step: 90, loss: 0.32485222816467285
step: 100, loss: 0.34782862663269043
step: 110, loss: 0.4785038232803345
step: 120, loss: 0.27841663360595703
step: 130, loss: 0.13115015625953674
step: 140, loss: 0.33052393794059753
step: 150, loss: 0.2578485310077667
step: 160, loss: 0.10112304240465164
step: 170, loss: 0.17520369589328766
step: 180, loss: 0.28176355361938477
step: 190, loss: 0.4980196952819824
step: 200, loss: 0.21258173882961273
step: 210, loss: 0.4097851514816284
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4163027656477438, f1=0.4604519774011299,
best_f1=0.4604519774011299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19420400261878967
step: 10, loss: 0.29102039337158203
step: 20, loss: 0.453408807516098
step: 30, loss: 0.29599666595458984
step: 40, loss: 0.10492093861103058
step: 50, loss: 0.3199425935745239
step: 60, loss: 0.07902472466230392
step: 70, loss: 0.35806676745414734
step: 80, loss: 0.1906835287809372
step: 90, loss: 0.28105705976486206
step: 100, loss: 0.05348317325115204
step: 110, loss: 0.1959485113620758
step: 120, loss: 0.13740362226963043
step: 130, loss: 0.4344751238822937
step: 140, loss: 0.39812928438186646
step: 150, loss: 0.11797311902046204
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.1216772049665451
step: 170, loss: 0.1205022782087326
step: 180, loss: 0.28551191091537476
step: 190, loss: 0.25391191244125366
step: 200, loss: 0.09522154182195663

```

```

step: 210, loss: 0.11896616220474243
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4480519480519481, f1=0.4936305732484077,
best_f1=0.4936305732484077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16911479830741882
step: 10, loss: 0.12378085404634476
step: 20, loss: 0.1695905178785324
step: 30, loss: 0.07460375130176544
step: 40, loss: 0.1536201387643814
step: 50, loss: 0.14760112762451172
step: 60, loss: 0.32510167360305786
step: 70, loss: 0.16059136390686035
step: 80, loss: 0.2951738238334656
step: 90, loss: 0.31483596563339233
step: 100, loss: 0.2559710741043091
step: 110, loss: 0.17980021238327026
step: 120, loss: 0.18514418601989746
step: 130, loss: 0.12945418059825897
step: 140, loss: 0.20625784993171692
step: 150, loss: 0.26778650283813477
step: 160, loss: 0.16572333872318268
step: 170, loss: 0.18966679275035858
step: 180, loss: 0.12279073894023895
step: 190, loss: 0.21135112643241882
step: 200, loss: 0.24025844037532806
step: 210, loss: 0.21285122632980347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.45307443365695793, f1=0.5047619047619047,
best_f1=0.5047619047619047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1478213667869568

```

```

step: 10, loss: 0.17322465777397156
step: 20, loss: 0.31569862365722656
step: 30, loss: 0.12820275127887726
step: 40, loss: 0.182456374168396
step: 50, loss: 0.14041005074977875
step: 60, loss: 0.22092635929584503
step: 70, loss: 0.09334752708673477
step: 80, loss: 0.1655760258436203
step: 90, loss: 0.04880460351705551
step: 100, loss: 0.12713396549224854
step: 110, loss: 0.18332473933696747
step: 120, loss: 0.08387583494186401
step: 130, loss: 0.17144080996513367
step: 140, loss: 0.1190577894449234
step: 150, loss: 0.19520175457000732
step: 160, loss: 0.07635627686977386
step: 170, loss: 0.12421946972608566
step: 180, loss: 0.1345585137605667
step: 190, loss: 0.3775899112224579
step: 200, loss: 0.12010009586811066
step: 210, loss: 0.21630039811134338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 14: dev_f1=0.46183206106870234, f1=0.5154264972776769,
best_f1=0.5154264972776769
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.046172965317964554
step: 10, loss: 0.15197470784187317
step: 20, loss: 0.32854506373405457
step: 30, loss: 0.08678928762674332
step: 40, loss: 0.24306301772594452
step: 50, loss: 0.0280526764690876
step: 60, loss: 0.28279808163642883
step: 70, loss: 0.2421526461839676
step: 80, loss: 0.17187076807022095
step: 90, loss: 0.24840673804283142
step: 100, loss: 0.22444573044776917
step: 110, loss: 0.12041544169187546
step: 120, loss: 0.18290762603282928
step: 130, loss: 0.15554386377334595
step: 140, loss: 0.16340190172195435

```

```

step: 150, loss: 0.18163855373859406
step: 160, loss: 0.3032277226448059
step: 170, loss: 0.24048668146133423
step: 180, loss: 0.17840923368930817
step: 190, loss: 0.19248506426811218
step: 200, loss: 0.14938265085220337
step: 210, loss: 0.31236928701400757
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.4692179700499168, f1=0.5039872408293461,
best_f1=0.5039872408293461
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:05, 448.50it/s]
load_f1 = 0.4656616415410385

```

```

real_f1 = 0.4542253521126761
267it [00:00, 1352.32it/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.")
4398it [00:21, 206.91it/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 roberta \
  --fp16 \
  --save_model \
  --dk general \
  --dk product \
  --summarize \
  --da entry_swap \
  --da attr_del

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

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

```

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

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.4413926303386688
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4805144965648651
step: 20, loss: 0.26865145564079285
step: 30, loss: 0.3644719421863556
step: 40, loss: 0.2436416894197464
step: 50, loss: 0.33831313252449036
step: 60, loss: 0.48728838562965393
step: 70, loss: 0.42347338795661926
step: 80, loss: 0.18110929429531097
step: 90, loss: 0.30010098218917847
step: 100, loss: 0.46501946449279785
step: 110, loss: 0.23515036702156067
step: 120, loss: 0.34642869234085083
step: 130, loss: 0.33230453729629517
step: 140, loss: 0.23083317279815674

```

```

step: 150, loss: 0.3090667724609375
step: 160, loss: 0.24019461870193481
step: 170, loss: 0.3976917266845703
step: 180, loss: 0.15657541155815125
step: 190, loss: 0.17245496809482574
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37430739402770996
step: 10, loss: 0.3016575872898102
step: 20, loss: 0.570421040058136
step: 30, loss: 0.2445693016052246
step: 40, loss: 0.5609502792358398
step: 50, loss: 0.3047938048839569
step: 60, loss: 0.4539155960083008
step: 70, loss: 0.3372689187526703
step: 80, loss: 0.16324910521507263
step: 90, loss: 0.3207288980484009
step: 100, loss: 0.26777365803718567
step: 110, loss: 0.3936818838119507
step: 120, loss: 0.23644253611564636
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 130, loss: 0.5109986066818237
step: 140, loss: 0.31828364729881287
step: 150, loss: 0.31073296070098877
step: 160, loss: 0.31615227460861206
step: 170, loss: 0.2402729094028473
step: 180, loss: 0.1693158596754074
step: 190, loss: 0.23781083524227142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 2: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.17216770740410348
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38175255060195923
step: 10, loss: 0.389522910118103

```

```

step: 20, loss: 0.4696469008922577
step: 30, loss: 0.29519981145858765
step: 40, loss: 0.08817644417285919
step: 50, loss: 0.38431045413017273
step: 60, loss: 0.1614934802055359
step: 70, loss: 0.3906395733356476
step: 80, loss: 0.30236271023750305
step: 90, loss: 0.38410353660583496
step: 100, loss: 0.5696932673454285
step: 110, loss: 0.6407153010368347
step: 120, loss: 0.3765340745449066
step: 130, loss: 0.1542743444442749
step: 140, loss: 0.34506574273109436
step: 150, loss: 0.3595145046710968
step: 160, loss: 0.6040363311767578
step: 170, loss: 0.4212009310722351
step: 180, loss: 0.35179802775382996
step: 190, loss: 0.1652681529521942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.2025040827436037, f1=0.20712328767123286,
best_f1=0.20712328767123286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2218940258026123
step: 10, loss: 0.21757349371910095
step: 20, loss: 0.2887326776981354
step: 30, loss: 0.24296733736991882
step: 40, loss: 0.5350006818771362
step: 50, loss: 0.23191122710704803
step: 60, loss: 0.3610233962535858
step: 70, loss: 0.2957395315170288
step: 80, loss: 0.22394779324531555
step: 90, loss: 0.15809538960456848
step: 100, loss: 0.38834407925605774
step: 110, loss: 0.4518400728702545
step: 120, loss: 0.2234535962343216
step: 130, loss: 0.417007178068161
step: 140, loss: 0.3862205743789673
step: 150, loss: 0.23154371976852417
step: 160, loss: 0.2905811071395874
step: 170, loss: 0.4059736132621765

```



```

step: 180, loss: 0.3571479618549347
step: 190, loss: 0.14057354629039764
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20378619153674837, f1=0.21176470588235297,
best_f1=0.21176470588235297
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.37923434376716614
step: 10, loss: 0.37021809816360474
step: 20, loss: 0.14582403004169464
step: 30, loss: 0.13055798411369324
step: 40, loss: 0.30362001061439514
step: 50, loss: 0.5514394640922546
step: 60, loss: 0.24009452760219574
step: 70, loss: 0.365360289812088
step: 80, loss: 0.34408918023109436
step: 90, loss: 0.2931373119354248
step: 100, loss: 0.4262986183166504
step: 110, loss: 0.3762189447879791
step: 120, loss: 0.22192630171775818
step: 130, loss: 0.5365946888923645
step: 140, loss: 0.37808847427368164
step: 150, loss: 0.2943319082260132
step: 160, loss: 0.15068358182907104
step: 170, loss: 0.3508046567440033
step: 180, loss: 0.22977793216705322
step: 190, loss: 0.2892818748950958
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20642201834862384, f1=0.21602787456445996,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29735323786735535
step: 10, loss: 0.23742976784706116

```

```

step: 20, loss: 0.280663400888443
step: 30, loss: 0.43941885232925415
step: 40, loss: 0.26666370034217834
step: 50, loss: 0.3009784519672394
step: 60, loss: 0.40883365273475647
step: 70, loss: 0.27119287848472595
step: 80, loss: 0.27377691864967346
step: 90, loss: 0.21695590019226074
step: 100, loss: 0.45225921273231506
step: 110, loss: 0.23539312183856964
step: 120, loss: 0.4526558816432953
step: 130, loss: 0.5246230363845825
step: 140, loss: 0.18006166815757751
step: 150, loss: 0.38379400968551636
step: 160, loss: 0.3828343152999878
step: 170, loss: 0.39236143231391907
step: 180, loss: 0.17870540916919708
step: 190, loss: 0.31160327792167664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.17216770740410348, f1=0.17216770740410348,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3142218589782715
step: 10, loss: 0.3072943389415741
step: 20, loss: 0.3052518665790558
step: 30, loss: 0.18731169402599335
step: 40, loss: 0.30792558193206787
step: 50, loss: 0.08631190657615662
step: 60, loss: 0.15559494495391846
step: 70, loss: 0.15790465474128723
step: 80, loss: 0.23475490510463715
step: 90, loss: 0.23500661551952362
step: 100, loss: 0.5797333717346191
step: 110, loss: 0.4321860671043396
step: 120, loss: 0.38392144441604614
step: 130, loss: 0.3108298182487488
step: 140, loss: 0.2385959029197693
step: 150, loss: 0.29428601264953613
step: 160, loss: 0.35843682289123535
step: 170, loss: 0.2708812355995178
step: 180, loss: 0.21560809016227722
step: 190, loss: 0.3052448332309723
/usr/local/lib/python3.7/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=0.20118983234180637, f1=0.20510037981551815,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2512564957141876
step: 10, loss: 0.37134429812431335
step: 20, loss: 0.2319032996892929
step: 30, loss: 0.3862839937210083
step: 40, loss: 0.15387076139450073
step: 50, loss: 0.3109617829322815
step: 60, loss: 0.434845894575119
step: 70, loss: 0.21801045536994934
step: 80, loss: 0.3599326014518738
step: 90, loss: 0.22286565601825714
step: 100, loss: 0.31399673223495483
step: 110, loss: 0.3786310851573944
step: 120, loss: 0.35214173793792725
step: 130, loss: 0.2866276800632477
step: 140, loss: 0.3627394437789917
step: 150, loss: 0.29704615473747253
step: 160, loss: 0.16631148755550385
step: 170, loss: 0.22174285352230072
step: 180, loss: 0.3019081950187683
step: 190, loss: 0.2751351296901703
/usr/local/lib/python3.7/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=0.20195439739413681, f1=0.20577027762656502,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22878757119178772
step: 10, loss: 0.09218954294919968
step: 20, loss: 0.29083752632141113
step: 30, loss: 0.17542031407356262
step: 40, loss: 0.30269595980644226
step: 50, loss: 0.355265736579895
step: 60, loss: 0.364860475063324
step: 70, loss: 0.16454458236694336
step: 80, loss: 0.3618741035461426
step: 90, loss: 0.7461548447608948
step: 100, loss: 0.3387950658798218
step: 110, loss: 0.3442023992538452
step: 120, loss: 0.5489190816879272
step: 130, loss: 0.2965348958969116

```

```

step: 140, loss: 0.30421724915504456
step: 150, loss: 0.23527655005455017
step: 160, loss: 0.21963052451610565
step: 170, loss: 0.518138587474823
step: 180, loss: 0.35277804732322693
step: 190, loss: 0.24505753815174103
/usr/local/lib/python3.7/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=0.2039911308203991, f1=0.20905533817775293,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28708773851394653
step: 10, loss: 0.1539716124534607
step: 20, loss: 0.28964823484420776
step: 30, loss: 0.3767397403717041
step: 40, loss: 0.28177374601364136
step: 50, loss: 0.09831574559211731
step: 60, loss: 0.304324209690094
step: 70, loss: 0.29745200276374817
step: 80, loss: 0.41691216826438904
step: 90, loss: 0.1535630077123642
step: 100, loss: 0.2173144817352295
step: 110, loss: 0.22118259966373444
step: 120, loss: 0.3461056351661682
step: 130, loss: 0.4723409116268158
step: 140, loss: 0.3419887125492096
step: 150, loss: 0.30455267429351807
step: 160, loss: 0.218044251203537
step: 170, loss: 0.3418397903442383
step: 180, loss: 0.22086156904697418
step: 190, loss: 0.15445150434970856
/usr/local/lib/python3.7/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=0.20378619153674837, f1=0.20840336134453782,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5098609328269958
step: 10, loss: 0.07829761505126953
step: 20, loss: 0.3167053759098053
step: 30, loss: 0.21654416620731354
step: 40, loss: 0.07739794999361038
step: 50, loss: 0.28905895352363586

```

```

step: 60, loss: 0.21828536689281464
step: 70, loss: 0.6122047901153564
step: 80, loss: 0.3730833828449249
step: 90, loss: 0.3931848108768463
step: 100, loss: 0.17832955718040466
step: 110, loss: 0.23694466054439545
step: 120, loss: 0.2881879210472107
step: 130, loss: 0.63383948802948
step: 140, loss: 0.4298315942287445
step: 150, loss: 0.30439266562461853
step: 160, loss: 0.299383282661438
step: 170, loss: 0.4579557180404663
step: 180, loss: 0.2822442352771759
step: 190, loss: 0.12251287698745728
/usr/local/lib/python3.7/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=0.2027700831024931, f1=0.2120866590649943,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3817349672317505
step: 10, loss: 0.3012619912624359
step: 20, loss: 0.5234796404838562
step: 30, loss: 0.3554191291332245
step: 40, loss: 0.2887103855609894
step: 50, loss: 0.47995731234550476
step: 60, loss: 0.4269101619720459
step: 70, loss: 0.3633294105529785
step: 80, loss: 0.3685402572154999
step: 90, loss: 0.4445873200893402
step: 100, loss: 0.28327664732933044
step: 110, loss: 0.5066455006599426
step: 120, loss: 0.16634756326675415
step: 130, loss: 0.2302316576242447
step: 140, loss: 0.3644101321697235
step: 150, loss: 0.34910136461257935
step: 160, loss: 0.22307369112968445
step: 170, loss: 0.4798341691493988
step: 180, loss: 0.28786352276802063
step: 190, loss: 0.1560572236776352
/usr/local/lib/python3.7/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=0.20408163265306123, f1=0.21485714285714286,
best_f1=0.21602787456445996
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2955910861492157
step: 10, loss: 0.28537318110466003
step: 20, loss: 0.35896170139312744
step: 30, loss: 0.3472120761871338
step: 40, loss: 0.2251691222190857
step: 50, loss: 0.23428891599178314
step: 60, loss: 0.2835833728313446
step: 70, loss: 0.15921303629875183
step: 80, loss: 0.2945389449596405
step: 90, loss: 0.2924869954586029
step: 100, loss: 0.1625908464193344
step: 110, loss: 0.5451594591140747
step: 120, loss: 0.1595245599746704
step: 130, loss: 0.22056256234645844
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 140, loss: 0.3025408685207367
step: 150, loss: 0.3554525375366211
step: 160, loss: 0.4459932744503021
step: 170, loss: 0.10622741281986237
step: 180, loss: 0.2262318879365921
step: 190, loss: 0.30817049741744995
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20670391061452514, f1=0.21339387060158913,
best_f1=0.21339387060158913
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28680655360221863
step: 10, loss: 0.22061468660831451
step: 20, loss: 0.3526844084262848
step: 30, loss: 0.10914671421051025
step: 40, loss: 0.30114006996154785
step: 50, loss: 0.3510630130767822
step: 60, loss: 0.2706139385700226
step: 70, loss: 0.23314666748046875
step: 80, loss: 0.36082327365875244
step: 90, loss: 0.13997118175029755
step: 100, loss: 0.2882431149482727
step: 110, loss: 0.49248477816581726
step: 120, loss: 0.28515389561653137

```

```

step: 130, loss: 0.2903381288051605
step: 140, loss: 0.16048182547092438
step: 150, loss: 0.4282926619052887
step: 160, loss: 0.39832165837287903
step: 170, loss: 0.20947343111038208
step: 180, loss: 0.29110050201416016
step: 190, loss: 0.5466832518577576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.20912124582869854, f1=0.21230942970073408,
best_f1=0.21230942970073408
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4981628656387329
step: 10, loss: 0.21815262734889984
step: 20, loss: 0.2310335785150528
step: 30, loss: 0.2890758216381073
step: 40, loss: 0.6357709765434265
step: 50, loss: 0.2950541079044342
step: 60, loss: 0.17066025733947754
step: 70, loss: 0.23292513191699982
step: 80, loss: 0.21201612055301666
step: 90, loss: 0.30699652433395386
step: 100, loss: 0.33367326855659485
step: 110, loss: 0.4111620783805847
step: 120, loss: 0.22575266659259796
step: 130, loss: 0.2850321829319
step: 140, loss: 0.5083796977996826
step: 150, loss: 0.13837766647338867
step: 160, loss: 0.28111588954925537
step: 170, loss: 0.7518206238746643
step: 180, loss: 0.28419333696365356
step: 190, loss: 0.413557767868042
/usr/local/lib/python3.7/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=0.20681818181818182, f1=0.21609195402298853,
best_f1=0.21230942970073408
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:08, 232.50it/s]
load_f1 = 0.20798201236649805
real_f1 = 0.20727673649393605
733it [00:00, 3433.49it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:21, 202.15it/s]

```

4.2 DITTO TEXTUAL

4.2.1 Abt-Buy - Running the matcher

```

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

```



```

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

```

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

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

Step 3: run the matcher

```

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

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly

```

because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5060041546821594
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.4880598783493042
step: 20, loss: 0.27633804082870483
step: 30, loss: 0.4530740976333618
step: 40, loss: 0.5353206992149353
step: 50, loss: 0.3098382353782654
step: 60, loss: 0.5936725735664368
step: 70, loss: 0.35039153695106506
step: 80, loss: 0.22264498472213745
step: 90, loss: 0.21479207277297974
step: 100, loss: 0.1643984317779541
step: 110, loss: 0.4461592733860016
step: 120, loss: 0.34587162733078003
step: 130, loss: 0.3611368238925934
step: 140, loss: 0.39522701501846313
step: 150, loss: 0.3057967722415924
step: 160, loss: 0.3990769386291504
step: 170, loss: 0.3344683349132538
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.1941564561734213, f1=0.1941564561734213,
best_f1=0.1941564561734213
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.33906692266464233
step: 10, loss: 0.4864880442619324

```

```

step: 20, loss: 0.3176894783973694
step: 30, loss: 0.3301257789134979
step: 40, loss: 0.07118536531925201
step: 50, loss: 0.470339834690094
step: 60, loss: 0.19974327087402344
step: 70, loss: 0.5068038702011108
step: 80, loss: 0.23098638653755188
step: 90, loss: 0.2421543002128601
step: 100, loss: 0.5191940665245056
step: 110, loss: 0.25779488682746887
step: 120, loss: 0.2362736016511917
step: 130, loss: 0.5618122816085815
step: 140, loss: 0.5066244602203369
step: 150, loss: 0.41780126094818115
step: 160, loss: 0.44054120779037476
step: 170, loss: 0.3534097373485565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.24251968503937008, f1=0.2396006655574043,
best_f1=0.2396006655574043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5625867247581482
step: 10, loss: 0.31542259454727173
step: 20, loss: 0.2557789981365204
step: 30, loss: 0.24818246066570282
step: 40, loss: 0.355487197637558
step: 50, loss: 0.5777066349983215
step: 60, loss: 0.2865064740180969
step: 70, loss: 0.24829308688640594
step: 80, loss: 0.378502756357193
step: 90, loss: 0.5336341857910156
step: 100, loss: 0.289023756980896
step: 110, loss: 0.1644231677055359
step: 120, loss: 0.5453325510025024
step: 130, loss: 0.483041375875473
step: 140, loss: 0.44631874561309814
step: 150, loss: 0.20220309495925903
step: 160, loss: 0.1583048403263092
step: 170, loss: 0.31653469800949097
/usr/local/lib/python3.7/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=0.23623445825932504, f1=0.23487544483985764,
best_f1=0.2396006655574043
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3915606439113617
step: 10, loss: 0.5392979383468628
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.22396422922611237
step: 30, loss: 0.3857465386390686
step: 40, loss: 0.2529415190219879
step: 50, loss: 0.3665289580821991
step: 60, loss: 0.6618731617927551
step: 70, loss: 0.31421172618865967
step: 80, loss: 0.4765564799308777
step: 90, loss: 0.28313636779785156
step: 100, loss: 0.3382607698440552
step: 110, loss: 0.4397784471511841
step: 120, loss: 0.47207164764404297
step: 130, loss: 0.33237650990486145
step: 140, loss: 0.22301506996154785
step: 150, loss: 0.7266283631324768
step: 160, loss: 0.14205439388751984
step: 170, loss: 0.23766571283340454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.26356589147286824, f1=0.2360034453057709,
best_f1=0.2360034453057709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3541010320186615
step: 10, loss: 0.29018154740333557
step: 20, loss: 0.28127333521842957
step: 30, loss: 0.21904537081718445
step: 40, loss: 0.2633052468299866
step: 50, loss: 0.2359689325094223
step: 60, loss: 0.3468209505081177
step: 70, loss: 0.336018830537796
step: 80, loss: 0.084278404712677
step: 90, loss: 0.5607019662857056
step: 100, loss: 0.24657496809959412

```

```

step: 110, loss: 0.25955766439437866
step: 120, loss: 0.15194636583328247
step: 130, loss: 0.2312816083431244
step: 140, loss: 0.19347137212753296
step: 150, loss: 0.20031997561454773
step: 160, loss: 0.2568568289279938
step: 170, loss: 0.3667233884334564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5649484536082475, f1=0.5766871165644172,
best_f1=0.5766871165644172
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26289063692092896
step: 10, loss: 0.10186257213354111
step: 20, loss: 0.35122203826904297
step: 30, loss: 0.1915929913520813
step: 40, loss: 0.3449437916278839
step: 50, loss: 0.1618650257587433
step: 60, loss: 0.21407780051231384
step: 70, loss: 0.1905411034822464
step: 80, loss: 0.21861372888088226
step: 90, loss: 0.16428956389427185
step: 100, loss: 0.17142271995544434
step: 110, loss: 0.281853049993515
step: 120, loss: 0.379944771528244
step: 130, loss: 0.3042697608470917
step: 140, loss: 0.2596501111984253
step: 150, loss: 0.20633552968502045
step: 160, loss: 0.2128293365240097
step: 170, loss: 0.11528918147087097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.6561797752808989,
best_f1=0.6561797752808989
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1229420080780983
step: 10, loss: 0.4333944618701935
step: 20, loss: 0.33469563722610474
step: 30, loss: 0.3667454123497009
step: 40, loss: 0.05341802537441254
step: 50, loss: 0.06032676249742508
step: 60, loss: 0.306702196598053
step: 70, loss: 0.12256386131048203
step: 80, loss: 0.09041278809309006
step: 90, loss: 0.14645737409591675
step: 100, loss: 0.11126208305358887
step: 110, loss: 0.16639341413974762
step: 120, loss: 0.09681151807308197
step: 130, loss: 0.29139018058776855
step: 140, loss: 0.06348942220211029
step: 150, loss: 0.1614767611026764
step: 160, loss: 0.17574194073677063
step: 170, loss: 0.19536617398262024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7155963302752293, f1=0.6894977168949772,
best_f1=0.6894977168949772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1413475126028061
step: 10, loss: 0.15384052693843842
step: 20, loss: 0.08192131668329239
step: 30, loss: 0.01324191503226757
step: 40, loss: 0.17788860201835632
step: 50, loss: 0.06446538120508194
step: 60, loss: 0.18429069221019745
step: 70, loss: 0.03332192450761795
step: 80, loss: 0.2670455873012543
step: 90, loss: 0.2191222906112671
step: 100, loss: 0.03763798251748085
step: 110, loss: 0.16640077531337738
step: 120, loss: 0.29674988985061646
step: 130, loss: 0.16050498187541962
step: 140, loss: 0.33749258518218994
step: 150, loss: 0.15778997540473938
step: 160, loss: 0.06282635778188705

```

```

step: 170, loss: 0.240537628531456
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7607655502392345, f1=0.722488038277512,
best_f1=0.722488038277512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1066632941365242
step: 10, loss: 0.03154850751161575
step: 20, loss: 0.10756921768188477
step: 30, loss: 0.16854019463062286
step: 40, loss: 0.15661756694316864
step: 50, loss: 0.02128683216869831
step: 60, loss: 0.12955830991268158
step: 70, loss: 0.31537124514579773
step: 80, loss: 0.09635628014802933
step: 90, loss: 0.162726029753685
step: 100, loss: 0.05995015427470207
step: 110, loss: 0.19593766331672668
step: 120, loss: 0.05526713281869888
step: 130, loss: 0.25423315167427063
step: 140, loss: 0.10882455855607986
step: 150, loss: 0.40828439593315125
step: 160, loss: 0.04244665056467056
step: 170, loss: 0.08900173008441925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7780678851174935, f1=0.7263157894736842,
best_f1=0.7263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03505707532167435
step: 10, loss: 0.04649568349123001
step: 20, loss: 0.1632087528705597
step: 30, loss: 0.1405094414949417
step: 40, loss: 0.11374890804290771

```

```

step: 50, loss: 0.11353342235088348
step: 60, loss: 0.11993279308080673
step: 70, loss: 0.20043043792247772
step: 80, loss: 0.17825865745544434
step: 90, loss: 0.1557127684354782
step: 100, loss: 0.09331152588129044
step: 110, loss: 0.11247048527002335
step: 120, loss: 0.1407402753829956
step: 130, loss: 0.04398823902010918
step: 140, loss: 0.08143622428178787
step: 150, loss: 0.11409679800271988
step: 160, loss: 0.18489333987236023
step: 170, loss: 0.17146334052085876
/usr/local/lib/python3.7/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=0.7642679900744417, f1=0.7427184466019418,
best_f1=0.7263157894736842
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1664874106645584
step: 10, loss: 0.022548140957951546
step: 20, loss: 0.04947103559970856
step: 30, loss: 0.056848201900720596
step: 40, loss: 0.014269369654357433
step: 50, loss: 0.07668165862560272
step: 60, loss: 0.14694824814796448
step: 70, loss: 0.02946287952363491
step: 80, loss: 0.0194704532623291
step: 90, loss: 0.253334641456604
step: 100, loss: 0.3157424032688141
step: 110, loss: 0.12055309861898422
step: 120, loss: 0.029350394383072853
step: 130, loss: 0.03994220122694969
step: 140, loss: 0.26475125551223755
step: 150, loss: 0.06032169982790947
step: 160, loss: 0.12491822242736816
step: 170, loss: 0.16680105030536652
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7830687830687831, f1=0.7455012853470437,
best_f1=0.7455012853470437

```


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

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

step: 0, loss: 0.018671540543437004
step: 10, loss: 0.06896831840276718
step: 20, loss: 0.06392174959182739
step: 30, loss: 0.10455796122550964
step: 40, loss: 0.019756287336349487
step: 50, loss: 0.04472819343209267
step: 60, loss: 0.05621194466948509
step: 70, loss: 0.19448956847190857
step: 80, loss: 0.040746744722127914
step: 90, loss: 0.035189978778362274
step: 100, loss: 0.018752263858914375
step: 110, loss: 0.0605769120156765
step: 120, loss: 0.08566481620073318
step: 130, loss: 0.18329095840454102
step: 140, loss: 0.020074833184480667
step: 150, loss: 0.041084084659814835
step: 160, loss: 0.10183383524417877
step: 170, loss: 0.363845556974411

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

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

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

warnings.warn(SAVE_STATE_WARNING, UserWarning)

epoch 12: dev_f1=0.7880434782608696, f1=0.7352185089974292,
best_f1=0.7352185089974292

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

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

step: 0, loss: 0.09657707065343857
step: 10, loss: 0.008610248565673828
step: 20, loss: 0.03078070655465126
step: 30, loss: 0.21418605744838715
step: 40, loss: 0.3043511211872101
step: 50, loss: 0.05669766664505005
step: 60, loss: 0.04021119698882103
step: 70, loss: 0.22928300499916077
step: 80, loss: 0.03957785665988922
step: 90, loss: 0.042749933898448944
step: 100, loss: 0.06330224871635437
step: 110, loss: 0.03237992525100708
step: 120, loss: 0.012598467990756035
step: 130, loss: 0.021083615720272064
step: 140, loss: 0.2873414158821106

```

step: 150, loss: 0.05469216778874397
step: 160, loss: 0.023392338305711746
step: 170, loss: 0.01070152036845684
/usr/local/lib/python3.7/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=0.783289817232376, f1=0.7444168734491315,
best_f1=0.7352185089974292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10256626456975937
step: 10, loss: 0.025750208646059036
step: 20, loss: 0.23494774103164673
step: 30, loss: 0.06872273981571198
step: 40, loss: 0.05069364234805107
step: 50, loss: 0.013733827508985996
step: 60, loss: 0.024672530591487885
step: 70, loss: 0.13627544045448303
step: 80, loss: 0.01744093932211399
step: 90, loss: 0.00846592616289854
step: 100, loss: 0.02581178955733776
step: 110, loss: 0.06061284616589546
step: 120, loss: 0.12326091527938843
step: 130, loss: 0.15177540481090546
step: 140, loss: 0.06041615456342697
step: 150, loss: 0.020301219075918198
step: 160, loss: 0.01893037185072899
step: 170, loss: 0.12213722616434097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7890818858560793, f1=0.7487922705314008,
best_f1=0.7487922705314008
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005648382939398289
step: 10, loss: 0.03031051903963089
step: 20, loss: 0.309005469083786
step: 30, loss: 0.20384754240512848
step: 40, loss: 0.03789255768060684
step: 50, loss: 0.0036531377118080854
step: 60, loss: 0.021601764485239983

```

```

step: 70, loss: 0.10686822980642319
step: 80, loss: 0.015000078827142715
step: 90, loss: 0.06504634767770767
step: 100, loss: 0.09096041321754456
step: 110, loss: 0.01298522762954235
step: 120, loss: 0.06818254292011261
step: 130, loss: 0.01393632311373949
step: 140, loss: 0.0321907103061676
step: 150, loss: 0.008761120028793812
step: 160, loss: 0.11861822009086609
step: 170, loss: 0.009669166058301926
/usr/local/lib/python3.7/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=0.7819548872180452, f1=0.7487684729064039,
best_f1=0.7487922705314008
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 313.71it/s]
load_f1 = 0.5777777777777778

```

```

real_f1 = 0.5382059800664453
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 196.52it/s]

```

4.3 DITTO DIRTY - oK

4.3.1 DBLP-ACM - Running the matcher

```

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

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

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

```

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

```

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5972806215286255
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 10, loss: 0.4554125666618347
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5248098373413086
step: 30, loss: 0.3351593315601349
step: 40, loss: 0.35310444235801697
step: 50, loss: 0.562762975692749
step: 60, loss: 0.4342520534992218
step: 70, loss: 0.29870346188545227
step: 80, loss: 0.2749083936214447
step: 90, loss: 0.15615807473659515
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.22101181745529175
step: 110, loss: 0.11631473153829575
step: 120, loss: 0.13535790145397186
step: 130, loss: 0.04352711886167526
step: 140, loss: 0.14842209219932556
step: 150, loss: 0.21008417010307312
step: 160, loss: 0.09486012160778046
step: 170, loss: 0.12346312403678894
step: 180, loss: 0.06370552629232407
step: 190, loss: 0.42113324999809265
step: 200, loss: 0.08265814930200577

```

```

step: 210, loss: 0.06023520231246948
step: 220, loss: 0.10516738891601562
step: 230, loss: 0.05229602009057999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9467969598262758, f1=0.943646408839779,
best_f1=0.943646408839779
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015852469950914383
step: 10, loss: 0.17107874155044556
step: 20, loss: 0.0797598659992218
step: 30, loss: 0.0694648027420044
step: 40, loss: 0.027219654992222786
step: 50, loss: 0.04339098185300827
step: 60, loss: 0.046431198716163635
step: 70, loss: 0.010068428702652454
step: 80, loss: 0.004072179086506367
step: 90, loss: 0.1322600543498993
step: 100, loss: 0.2337336540222168
step: 110, loss: 0.02892232872545719
step: 120, loss: 0.1214178055524826
step: 130, loss: 0.0720064789056778
step: 140, loss: 0.06742073595523834
step: 150, loss: 0.18896037340164185
step: 160, loss: 0.018919581547379494
step: 170, loss: 0.003982663620263338
step: 180, loss: 0.061565447598695755
step: 190, loss: 0.02862936072051525
step: 200, loss: 0.11687719076871872
step: 210, loss: 0.02745508961379528
step: 220, loss: 0.013424213044345379
step: 230, loss: 0.007205005269497633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9537444933920706, f1=0.9481981981981982,
best_f1=0.9481981981981982

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04484980180859566
step: 10, loss: 0.0801299586892128
step: 20, loss: 0.033526115119457245
step: 30, loss: 0.07204493880271912
step: 40, loss: 0.03787604346871376
step: 50, loss: 0.05660300329327583
step: 60, loss: 0.06313180178403854
step: 70, loss: 0.030764926224946976
step: 80, loss: 0.08597508817911148
step: 90, loss: 0.011023773811757565
step: 100, loss: 0.023636171594262123
step: 110, loss: 0.007316816132515669
step: 120, loss: 0.007604497950524092
step: 130, loss: 0.015572786331176758
step: 140, loss: 0.009905675426125526
step: 150, loss: 0.049915749579668045
step: 160, loss: 0.015274831093847752
step: 170, loss: 0.009178652428090572
step: 180, loss: 0.020248400047421455
step: 190, loss: 0.007997553795576096
step: 200, loss: 0.012982726097106934
step: 210, loss: 0.0026150685735046864
step: 220, loss: 0.04310913011431694
step: 230, loss: 0.05444706976413727
/usr/local/lib/python3.7/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=0.9511918274687856, f1=0.9457900807381776,
best_f1=0.9481981981981982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023201853036880493
step: 10, loss: 0.017051320523023605
step: 20, loss: 0.002251823665574193
step: 30, loss: 0.0037937758024781942
step: 40, loss: 0.14700503647327423
step: 50, loss: 0.013353331945836544
step: 60, loss: 0.014673782512545586
step: 70, loss: 0.10764755308628082
step: 80, loss: 0.11858736723661423
step: 90, loss: 0.021219471469521523
step: 100, loss: 0.003529746551066637
step: 110, loss: 0.003484027925878763
step: 120, loss: 0.11283724009990692

```

```

step: 130, loss: 0.08928230404853821
step: 140, loss: 0.008589657954871655
step: 150, loss: 0.02713160216808319
step: 160, loss: 0.06640433520078659
step: 170, loss: 0.025230158120393753
step: 180, loss: 0.1645781248807907
step: 190, loss: 0.00574903842061758
step: 200, loss: 0.03403308987617493
step: 210, loss: 0.01709354855120182
step: 220, loss: 0.0010800431482493877
step: 230, loss: 0.0037471929099410772
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9594742606790799, f1=0.9493813273340832,
best_f1=0.9493813273340832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002146544400602579
step: 10, loss: 0.01573726162314415
step: 20, loss: 0.20292268693447113
step: 30, loss: 0.005249124951660633
step: 40, loss: 0.0049317749217152596
step: 50, loss: 0.004551073536276817
step: 60, loss: 0.09924184530973434
step: 70, loss: 0.020289242267608643
step: 80, loss: 0.018154725432395935
step: 90, loss: 0.10495663434267044
step: 100, loss: 0.007626193575561047
step: 110, loss: 0.0009455239633098245
step: 120, loss: 0.15179036557674408
step: 130, loss: 0.04217430576682091
step: 140, loss: 0.06058280169963837
step: 150, loss: 0.0067688473500311375
step: 160, loss: 0.0017841706285253167
step: 170, loss: 0.026576131582260132
step: 180, loss: 0.005338519345968962
step: 190, loss: 0.034145209938287735
step: 200, loss: 0.04190930351614952
step: 210, loss: 0.02011852152645588
step: 220, loss: 0.007086853962391615
step: 230, loss: 0.030569054186344147
/usr/local/lib/python3.7/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=0.9507119386637459, f1=0.9439461883408071,
best_f1=0.9493813273340832
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002811812562867999
step: 10, loss: 0.022166326642036438
step: 20, loss: 0.003552391892299056
step: 30, loss: 0.002602888736873865
step: 40, loss: 0.0006794155924580991
step: 50, loss: 0.012333580292761326
step: 60, loss: 0.09211872518062592
step: 70, loss: 0.10556045174598694
step: 80, loss: 0.007049995008856058
step: 90, loss: 0.06229085847735405
step: 100, loss: 0.0011497262166813016
step: 110, loss: 0.03331795707345009
step: 120, loss: 0.0035140984691679478
step: 130, loss: 0.0037546788807958364
step: 140, loss: 0.0011580846039578319
step: 150, loss: 0.0008483560523018241
step: 160, loss: 0.0022581988014280796
step: 170, loss: 0.0006557874730788171
step: 180, loss: 0.030423898249864578
step: 190, loss: 0.003678077831864357
step: 200, loss: 0.167327418923378
step: 210, loss: 0.00891948863863945
step: 220, loss: 0.04202048480510712
step: 230, loss: 0.0031306983437389135
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9648924122310306, f1=0.9519450800915331,
best_f1=0.9519450800915331
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006519312039017677
step: 10, loss: 0.002227778546512127
step: 20, loss: 0.0007307514897547662
step: 30, loss: 0.0005470510805025697
step: 40, loss: 0.000887205358594656

```

```

step: 50, loss: 0.0062931375578045845
step: 60, loss: 0.0017369026318192482
step: 70, loss: 0.0019966335967183113
step: 80, loss: 0.0012712046736851335
step: 90, loss: 0.0016881362535059452
step: 100, loss: 0.0006344317807815969
step: 110, loss: 0.0013738642446696758
step: 120, loss: 0.0024711876176297665
step: 130, loss: 0.03907981887459755
step: 140, loss: 0.0024668416008353233
step: 150, loss: 0.1244373619556427
step: 160, loss: 0.0007007477688603103
step: 170, loss: 0.0548025406897068
step: 180, loss: 0.002372225746512413
step: 190, loss: 0.1534644067287445
step: 200, loss: 0.031218089163303375
step: 210, loss: 0.019186055287718773
step: 220, loss: 0.003796808421611786
step: 230, loss: 0.0018305518897250295
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.968609865470852, f1=0.9567198177676537,
best_f1=0.9567198177676537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007072238950058818
step: 10, loss: 0.023957695811986923
step: 20, loss: 0.0018532982794567943
step: 30, loss: 0.0007595912902615964
step: 40, loss: 0.0060880049131810665
step: 50, loss: 0.0008024809067137539
step: 60, loss: 0.0008241672767326236
step: 70, loss: 0.005593892186880112
step: 80, loss: 0.0024506570771336555
step: 90, loss: 0.000645727850496769
step: 100, loss: 0.0016090720891952515
step: 110, loss: 0.030996689572930336
step: 120, loss: 0.00036232025013305247
step: 130, loss: 0.034062597900629044
step: 140, loss: 0.0014617738779634237
step: 150, loss: 0.2659231424331665
step: 160, loss: 0.002697056857869029

```

```

step: 170, loss: 0.01428176835179329
step: 180, loss: 0.002373988041654229
step: 190, loss: 0.03680672124028206
step: 200, loss: 0.011325489729642868
step: 210, loss: 0.009013655595481396
step: 220, loss: 0.007401086390018463
step: 230, loss: 0.0027922876179218292
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9775280898876404, f1=0.9558323895809739,
best_f1=0.9558323895809739
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017279449384659529
step: 10, loss: 0.001193289877846837
step: 20, loss: 0.00443172687664628
step: 30, loss: 0.00106845295522362
step: 40, loss: 0.003911031410098076
step: 50, loss: 0.0010200702818110585
step: 60, loss: 0.0006660491926595569
step: 70, loss: 0.03224031999707222
step: 80, loss: 0.002191763138398528
step: 90, loss: 0.1248253583908081
step: 100, loss: 0.0004569469019770622
step: 110, loss: 0.000456102192401886
step: 120, loss: 0.009236113168299198
step: 130, loss: 0.06179096922278404
step: 140, loss: 0.050708796828985214
step: 150, loss: 0.01588594913482666
step: 160, loss: 0.009270149283111095
step: 170, loss: 0.0010469520930200815
step: 180, loss: 0.0013047928223386407
step: 190, loss: 0.00015724723925814033
step: 200, loss: 0.0045041427947580814
step: 210, loss: 0.0008097795071080327
step: 220, loss: 0.0004351649549789727
step: 230, loss: 0.00022701820125803351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9807474518686297, f1=0.9540229885057471,
best_f1=0.9540229885057471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000551258388441056
step: 10, loss: 0.0007342266035266221
step: 20, loss: 0.0004012267745565623
step: 30, loss: 0.00021580937027465552
step: 40, loss: 0.02714705467224121
step: 50, loss: 0.00012175818847026676
step: 60, loss: 0.0006544762873090804
step: 70, loss: 0.010140052996575832
step: 80, loss: 0.003828460816293955
step: 90, loss: 0.007090689614415169
step: 100, loss: 0.00037309175240807235
step: 110, loss: 0.011275336146354675
step: 120, loss: 0.00039247979293577373
step: 130, loss: 0.004659566562622786
step: 140, loss: 0.00021647359244525433
step: 150, loss: 0.06027532368898392
step: 160, loss: 0.0002071074559353292
step: 170, loss: 0.0001946765260072425
step: 180, loss: 0.001274285139515996
step: 190, loss: 0.0005375199834816158
step: 200, loss: 0.000593475706409663
step: 210, loss: 0.0014662526082247496
step: 220, loss: 0.0011343626538291574
step: 230, loss: 0.0006577984895557165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9821029082774049, f1=0.962962962962963,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014949965989217162
step: 10, loss: 0.0009044166654348373
step: 20, loss: 0.001158138969913125
step: 30, loss: 0.0004071598523296416
step: 40, loss: 0.00011358941264916211

```

```
step: 50, loss: 0.0007289405330084264
step: 60, loss: 0.001758136204443872
step: 70, loss: 0.0003888123028445989
step: 80, loss: 0.0016928269760683179
step: 90, loss: 0.22079820930957794
step: 100, loss: 0.0007576961652375758
step: 110, loss: 0.0003939488378819078
step: 120, loss: 0.0013944566017016768
step: 130, loss: 0.0003101684560533613
step: 140, loss: 0.001384776085615158
step: 150, loss: 0.0006322851404547691
step: 160, loss: 0.0009026431362144649
step: 170, loss: 0.003300813026726246
step: 180, loss: 0.0004548478464130312
step: 190, loss: 0.0005787698901258409
step: 200, loss: 0.003744088113307953
step: 210, loss: 0.00034888743539340794
step: 220, loss: 0.0004694641684181988
step: 230, loss: 0.0007422319613397121
/usr/local/lib/python3.7/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=0.978675645342312, f1=0.963882618510158,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005826579872518778
step: 10, loss: 0.00025363065651617944
step: 20, loss: 0.0011427890276536345
step: 30, loss: 0.002957334741950035
step: 40, loss: 0.037106528878211975
step: 50, loss: 0.0007531827432103455
step: 60, loss: 0.0007013165741227567
step: 70, loss: 0.0006650614668615162
step: 80, loss: 0.00010826090874616057
step: 90, loss: 0.001531588495709002
step: 100, loss: 0.00010398480662843212
step: 110, loss: 0.0011114967055618763
step: 120, loss: 0.003264122176915407
step: 130, loss: 0.00030788721051067114
step: 140, loss: 0.00065907760290429
step: 150, loss: 0.0009456360712647438
step: 160, loss: 0.001595564535818994
step: 170, loss: 0.0010870574042201042
step: 180, loss: 0.0002393030736129731
step: 190, loss: 0.000952318892814219
step: 200, loss: 0.0002809930592775345
```

```

step: 210, loss: 0.08111440390348434
step: 220, loss: 0.0014133262448012829
step: 230, loss: 0.0009357769740745425
/usr/local/lib/python3.7/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=0.9808773903262092, f1=0.9659090909090909,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008368974085897207
step: 10, loss: 0.0005369903519749641
step: 20, loss: 0.0005642111063934863
step: 30, loss: 0.0004599870298989117
step: 40, loss: 0.001664165174588561
step: 50, loss: 0.0068158432841300964
step: 60, loss: 0.000329061527736485
step: 70, loss: 0.0017554913647472858
step: 80, loss: 0.007073315791785717
step: 90, loss: 0.0002172800013795495
step: 100, loss: 0.02316305600106716
step: 110, loss: 0.012495378032326698
step: 120, loss: 0.002622873056679964
step: 130, loss: 0.0014092018827795982
step: 140, loss: 0.0001811763650039211
step: 150, loss: 0.0022975183092057705
step: 160, loss: 0.0015676531475037336
step: 170, loss: 0.0007082029478624463
step: 180, loss: 0.03766029700636864
step: 190, loss: 0.0008156529511325061
step: 200, loss: 0.0004624048597179353
step: 210, loss: 0.0004143610130995512
step: 220, loss: 0.0004450095584616065
step: 230, loss: 0.00033535680267959833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 13: dev_f1=0.9842696629213483, f1=0.963718820861678,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005393826868385077

```

```

step: 10, loss: 0.002225077012553811
step: 20, loss: 0.0015857235994189978
step: 30, loss: 0.00047764467308297753
step: 40, loss: 0.0015182839706540108
step: 50, loss: 0.00015699518553446978
step: 60, loss: 0.00022875559807289392
step: 70, loss: 0.0004303519381210208
step: 80, loss: 0.00023427425185218453
step: 90, loss: 0.0004494866298045963
step: 100, loss: 0.000599250546656549
step: 110, loss: 0.0004988994915038347
step: 120, loss: 0.00018544081831350923
step: 130, loss: 0.0008301000925712287
step: 140, loss: 0.0721735879778862
step: 150, loss: 0.0001715696562314406
step: 160, loss: 0.0004934826865792274
step: 170, loss: 0.00034682286786846817
step: 180, loss: 0.0003182046639267355
step: 190, loss: 0.000548427109606564
step: 200, loss: 0.0004805278149433434
step: 210, loss: 0.0017333622090518475
step: 220, loss: 0.001915041240863502
step: 230, loss: 0.03151898458600044
/usr/local/lib/python3.7/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=0.9831271091113611, f1=0.9648127128263336,
best_f1=0.963718820861678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007690976839512587
step: 10, loss: 0.00041243276791647077
step: 20, loss: 0.0005478262901306152
step: 30, loss: 0.005552513990551233
step: 40, loss: 0.0008088489412330091
step: 50, loss: 0.0002834867627825588
step: 60, loss: 0.0003432536032050848
step: 70, loss: 0.005195676814764738
step: 80, loss: 0.00019390102534089237
step: 90, loss: 0.0007755299448035657
step: 100, loss: 0.002522716298699379
step: 110, loss: 0.00031732494244351983
step: 120, loss: 0.035635825246572495
step: 130, loss: 0.0002746782556641847
step: 140, loss: 0.0010216940427199006
step: 150, loss: 0.0004679904959630221
step: 160, loss: 0.004393601790070534

```

```

step: 170, loss: 0.00021513823594432324
step: 180, loss: 0.0006345478468574584
step: 190, loss: 0.00044253820669837296
step: 200, loss: 0.008088800124824047
step: 210, loss: 0.006295075174421072
step: 220, loss: 0.0015342560363933444
step: 230, loss: 0.0010856003500521183
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9831649831649831, f1=0.9625425652667423,
best_f1=0.963718820861678
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 207.79it/s]
load_f1 = 0.9808773903262092
real_f1 = 0.9766925638179801
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```


4398it [00:22, 197.59it/s]

4.3.2 DBLP-GoogleScholar - Running the matcher

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

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

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

Downloading: 100% 481/481 [00:00<00:00, 430kB/s]
Downloading: 100% 899k/899k [00:01<00:00, 792kB/s]
Downloading: 100% 456k/456k [00:01<00:00, 401kB/s]
Downloading: 100% 501M/501M [00:07<00:00, 64.7MB/s]
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6149956583976746
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.49291667342185974
step: 20, loss: 0.31033945083618164
step: 30, loss: 0.3375636041164398
step: 40, loss: 0.36393681168556213
step: 50, loss: 0.6953979134559631
step: 60, loss: 0.35136836767196655
step: 70, loss: 0.4533482491970062
step: 80, loss: 0.4124554991722107
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
step: 90, loss: 0.5335617661476135
step: 100, loss: 0.5473364591598511
step: 110, loss: 0.19425559043884277
step: 120, loss: 0.20914484560489655
step: 130, loss: 0.08789416402578354
step: 140, loss: 0.26324889063835144
step: 150, loss: 0.20239906013011932
step: 160, loss: 0.21111711859703064
step: 170, loss: 0.2714550495147705
step: 180, loss: 0.1836307793855667
step: 190, loss: 0.2662494480609894
step: 200, loss: 0.2575524151325226
step: 210, loss: 0.1345183551311493

```

```

step: 220, loss: 0.14470162987709045
step: 230, loss: 0.2609466314315796
step: 240, loss: 0.08364308625459671
step: 250, loss: 0.09783579409122467
step: 260, loss: 0.48507264256477356
step: 270, loss: 0.27932748198509216
step: 280, loss: 0.06812874972820282
step: 290, loss: 0.06668832898139954
step: 300, loss: 0.11307282745838165
step: 310, loss: 0.21135570108890533
step: 320, loss: 0.1882483959197998
step: 330, loss: 0.06027834489941597
step: 340, loss: 0.4862723648548126
step: 350, loss: 0.2368806153535843
step: 360, loss: 0.006419147364795208
step: 370, loss: 0.03125012293457985
step: 380, loss: 0.09819621592760086
step: 390, loss: 0.0549364909529686
step: 400, loss: 0.10107485204935074
step: 410, loss: 0.21242879331111908
step: 420, loss: 0.03029101900756359
step: 430, loss: 0.04500589519739151
step: 440, loss: 0.08048424869775772
step: 450, loss: 0.23590770363807678
step: 460, loss: 0.07231741398572922
step: 470, loss: 0.07369717955589294
step: 480, loss: 0.17013248801231384
step: 490, loss: 0.11939725279808044
step: 500, loss: 0.047038041055202484
step: 510, loss: 0.08141371607780457
step: 520, loss: 0.10015100240707397
step: 530, loss: 0.08539576083421707
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8955091151622944, f1=0.8939802336028752,
best_f1=0.8939802336028752
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1007990762591362
step: 10, loss: 0.1627860963344574
step: 20, loss: 0.15409424901008606
step: 30, loss: 0.10433557629585266

```

step: 40, loss: 0.17907372117042542
step: 50, loss: 0.1050684005022049
step: 60, loss: 0.2274089902639389
step: 70, loss: 0.06585751473903656
step: 80, loss: 0.06719812750816345
step: 90, loss: 0.037091489881277084
step: 100, loss: 0.09025919437408447
step: 110, loss: 0.06181533634662628
step: 120, loss: 0.1289239227771759
step: 130, loss: 0.03606080636382103
step: 140, loss: 0.06660792976617813
step: 150, loss: 0.10509931296110153
step: 160, loss: 0.04237022623419762
step: 170, loss: 0.07856335490942001
step: 180, loss: 0.1273140013217926
step: 190, loss: 0.06185116246342659
step: 200, loss: 0.24449101090431213
step: 210, loss: 0.06823339313268661
step: 220, loss: 0.008547027595341206
step: 230, loss: 0.15507742762565613
step: 240, loss: 0.02130226604640484
step: 250, loss: 0.07968943566083908
step: 260, loss: 0.09414280951023102
step: 270, loss: 0.02403377927839756
step: 280, loss: 0.09982222318649292
step: 290, loss: 0.026100266724824905
step: 300, loss: 0.09005782008171082
step: 310, loss: 0.056780632585287094
step: 320, loss: 0.06981226801872253
step: 330, loss: 0.10006522387266159
step: 340, loss: 0.15894697606563568
step: 350, loss: 0.0066567156463861465
step: 360, loss: 0.11703865230083466
step: 370, loss: 0.08979804813861847
step: 380, loss: 0.12046308070421219
step: 390, loss: 0.013469541445374489
step: 400, loss: 0.16047430038452148
step: 410, loss: 0.10487312078475952
step: 420, loss: 0.05280162766575813
step: 430, loss: 0.07275494188070297
step: 440, loss: 0.03126987814903259
step: 450, loss: 0.05842221900820732
step: 460, loss: 0.036339469254016876
step: 470, loss: 0.06408408284187317
step: 480, loss: 0.01548478938639164
step: 490, loss: 0.09476468712091446
step: 500, loss: 0.024283891543745995
step: 510, loss: 0.04683101922273636

```

step: 520, loss: 0.484140008687973
step: 530, loss: 0.08369224518537521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9294492489758761, f1=0.9216757741347906,
best_f1=0.9216757741347906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16363120079040527
step: 10, loss: 0.1373404562473297
step: 20, loss: 0.027972718700766563
step: 30, loss: 0.05122220516204834
step: 40, loss: 0.04271459951996803
step: 50, loss: 0.09446607530117035
step: 60, loss: 0.058138709515333176
step: 70, loss: 0.03253115341067314
step: 80, loss: 0.1846950650215149
step: 90, loss: 0.02808123268187046
step: 100, loss: 0.13442876935005188
step: 110, loss: 0.046069879084825516
step: 120, loss: 0.07992478460073471
step: 130, loss: 0.10410935431718826
step: 140, loss: 0.09330940246582031
step: 150, loss: 0.0704571008682251
step: 160, loss: 0.08973973244428635
step: 170, loss: 0.025375567376613617
step: 180, loss: 0.015905600041151047
step: 190, loss: 0.016476569697260857
step: 200, loss: 0.008463834412395954
step: 210, loss: 0.07018928974866867
step: 220, loss: 0.12988807260990143
step: 230, loss: 0.0332605354487896
step: 240, loss: 0.06375043094158173
step: 250, loss: 0.22071102261543274
step: 260, loss: 0.07848740369081497
step: 270, loss: 0.0643150806427002
step: 280, loss: 0.015764568001031876
step: 290, loss: 0.018997257575392723
step: 300, loss: 0.11808480322360992
step: 310, loss: 0.10841973125934601
step: 320, loss: 0.04124142974615097
step: 330, loss: 0.012318400666117668

```

```

step: 340, loss: 0.023597804829478264
step: 350, loss: 0.13916915655136108
step: 360, loss: 0.013406039215624332
step: 370, loss: 0.039764851331710815
step: 380, loss: 0.009757510386407375
step: 390, loss: 0.06296615302562714
step: 400, loss: 0.05210699513554573
step: 410, loss: 0.012043111957609653
step: 420, loss: 0.00969256367534399
step: 430, loss: 0.06052885949611664
step: 440, loss: 0.2900448739528656
step: 450, loss: 0.15378926694393158
step: 460, loss: 0.12450769543647766
step: 470, loss: 0.012566924095153809
step: 480, loss: 0.17602433264255524
step: 490, loss: 0.006511949468404055
step: 500, loss: 0.021809810772538185
step: 510, loss: 0.02495436556637287
step: 520, loss: 0.08510202169418335
step: 530, loss: 0.026674555614590645
/usr/local/lib/python3.7/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=0.925497454881999, f1=0.9178208679593721,
best_f1=0.9216757741347906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013723472133278847
step: 10, loss: 0.011738480068743229
step: 20, loss: 0.017798764631152153
step: 30, loss: 0.12538231909275055
step: 40, loss: 0.0524861142039299
step: 50, loss: 0.05753777548670769
step: 60, loss: 0.04041073098778725
step: 70, loss: 0.15069158375263214
step: 80, loss: 0.15584328770637512
step: 90, loss: 0.03652753308415413
step: 100, loss: 0.010986795648932457
step: 110, loss: 0.18475604057312012
step: 120, loss: 0.026720890775322914
step: 130, loss: 0.18255537748336792
step: 140, loss: 0.04978523775935173
step: 150, loss: 0.0409136638045311
step: 160, loss: 0.04667404666543007
step: 170, loss: 0.061558451503515244
step: 180, loss: 0.13289305567741394
step: 190, loss: 0.16363807022571564

```

```

step: 200, loss: 0.02770746499300003
step: 210, loss: 0.006588564720004797
step: 220, loss: 0.014420682564377785
step: 230, loss: 0.05157316476106644
step: 240, loss: 0.06549005210399628
step: 250, loss: 0.15794317424297333
step: 260, loss: 0.042120471596717834
step: 270, loss: 0.057284072041511536
step: 280, loss: 0.0035236021503806114
step: 290, loss: 0.028136800974607468
step: 300, loss: 0.003563364502042532
step: 310, loss: 0.006643139757215977
step: 320, loss: 0.15866853296756744
step: 330, loss: 0.026683256030082703
step: 340, loss: 0.03749867156147957
step: 350, loss: 0.01993916556239128
step: 360, loss: 0.028545618057250977
step: 370, loss: 0.060275640338659286
step: 380, loss: 0.06249735504388809
step: 390, loss: 0.007875756360590458
step: 400, loss: 0.04414300620555878
step: 410, loss: 0.023572329431772232
step: 420, loss: 0.015318972989916801
step: 430, loss: 0.0191352479159832
step: 440, loss: 0.0140862250700593
step: 450, loss: 0.08685409277677536
step: 460, loss: 0.03463626652956009
step: 470, loss: 0.008673489093780518
step: 480, loss: 0.06840655207633972
step: 490, loss: 0.0006272522732615471
step: 500, loss: 0.07902742922306061
step: 510, loss: 0.0531180277466774
step: 520, loss: 0.01423724740743637
step: 530, loss: 0.08290515840053558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9327731092436975, f1=0.9187411930483794,
best_f1=0.9187411930483794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031166696920990944
step: 10, loss: 0.010591219179332256

```

step: 20, loss: 0.024683183059096336
step: 30, loss: 0.034465380012989044
step: 40, loss: 0.06110921502113342
step: 50, loss: 0.07581395655870438
step: 60, loss: 0.022242307662963867
step: 70, loss: 0.08546223491430283
step: 80, loss: 0.002636958612129092
step: 90, loss: 0.15147435665130615
step: 100, loss: 0.1684953272342682
step: 110, loss: 0.016494493931531906
step: 120, loss: 0.015294533222913742
step: 130, loss: 0.011376520618796349
step: 140, loss: 0.060894764959812164
step: 150, loss: 0.038217898458242416
step: 160, loss: 0.027914853766560555
step: 170, loss: 0.14045719802379608
step: 180, loss: 0.01835952140390873
step: 190, loss: 0.004067596048116684
step: 200, loss: 0.010020517744123936
step: 210, loss: 0.0035483406390994787
step: 220, loss: 0.0049332138150930405
step: 230, loss: 0.0035547546576708555
step: 240, loss: 0.020203515887260437
step: 250, loss: 0.09366510063409805
step: 260, loss: 0.003417791798710823
step: 270, loss: 0.00865147728472948
step: 280, loss: 0.009847649373114109
step: 290, loss: 0.012597513385117054
step: 300, loss: 0.3683679699897766
step: 310, loss: 0.0653829425573349
step: 320, loss: 0.08218821883201599
step: 330, loss: 0.03001718968153
step: 340, loss: 0.014591231010854244
step: 350, loss: 0.0028321545105427504
step: 360, loss: 0.007088627200573683
step: 370, loss: 0.028093371540308
step: 380, loss: 0.02528325468301773
step: 390, loss: 0.025568583980202675
step: 400, loss: 0.03160492330789566
step: 410, loss: 0.04111652076244354
step: 420, loss: 0.1750514656305313
step: 430, loss: 0.020621811971068382
step: 440, loss: 0.006482795812189579
step: 450, loss: 0.01434569526463747
step: 460, loss: 0.18724249303340912
step: 470, loss: 0.07377646118402481
step: 480, loss: 0.03545133024454117
step: 490, loss: 0.017069755122065544


```

step: 500, loss: 0.046908408403396606
step: 510, loss: 0.017380358651280403
step: 520, loss: 0.20962633192539215
step: 530, loss: 0.06119368597865105
/usr/local/lib/python3.7/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=0.9319129226493746, f1=0.9187411930483794,
best_f1=0.9187411930483794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061247166246175766
step: 10, loss: 0.02046976424753666
step: 20, loss: 0.019171562045812607
step: 30, loss: 0.0008161202422343194
step: 40, loss: 0.0008995727985166013
step: 50, loss: 0.005951219704002142
step: 60, loss: 0.007993313483893871
step: 70, loss: 0.002945552347227931
step: 80, loss: 0.019846521317958832
step: 90, loss: 0.16896434128284454
step: 100, loss: 0.0304731335490942
step: 110, loss: 0.012163816019892693
step: 120, loss: 0.02763316221535206
step: 130, loss: 0.006606435868889093
step: 140, loss: 0.0066270302049815655
step: 150, loss: 0.0013900543563067913
step: 160, loss: 0.01976541429758072
step: 170, loss: 0.004899074323475361
step: 180, loss: 0.001443265238776803
step: 190, loss: 0.3627907931804657
step: 200, loss: 0.07142496854066849
step: 210, loss: 0.014901591464877129
step: 220, loss: 0.06860989332199097
step: 230, loss: 0.026612555608153343
step: 240, loss: 0.005356178618967533
step: 250, loss: 0.04792769253253937
step: 260, loss: 0.01973000355064869
step: 270, loss: 0.004422212950885296
step: 280, loss: 0.012046754360198975
step: 290, loss: 0.0018208562396466732
step: 300, loss: 0.0019116874318569899
step: 310, loss: 0.12910524010658264
step: 320, loss: 0.007166722323745489
step: 330, loss: 0.030379561707377434
step: 340, loss: 0.001516504562459886
step: 350, loss: 0.06036382541060448

```

```
step: 360, loss: 0.1407424360513687
step: 370, loss: 0.01253095269203186
step: 380, loss: 0.0005848266300745308
step: 390, loss: 0.04207876697182655
step: 400, loss: 0.003076678840443492
step: 410, loss: 0.04541020840406418
step: 420, loss: 0.00890765804797411
step: 430, loss: 0.005269213113933802
step: 440, loss: 0.003505873493850231
step: 450, loss: 0.2072111964225769
step: 460, loss: 0.006710894405841827
step: 470, loss: 0.01693830080330372
step: 480, loss: 0.014783350750803947
step: 490, loss: 0.006568441167473793
step: 500, loss: 0.022063355892896652
step: 510, loss: 0.010129219852387905
step: 520, loss: 0.005002585239708424
step: 530, loss: 0.005949050188064575
/usr/local/lib/python3.7/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=0.9259962049335864, f1=0.9153515064562409,
best_f1=0.9187411930483794
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0073462361469864845
step: 10, loss: 0.01571454294025898
step: 20, loss: 0.036157265305519104
step: 30, loss: 0.03754078224301338
step: 40, loss: 0.00669083371758461
step: 50, loss: 0.15207615494728088
step: 60, loss: 0.04943373426795006
step: 70, loss: 0.0018797715893015265
step: 80, loss: 0.0033200865145772696
step: 90, loss: 0.01609819382429123
step: 100, loss: 0.014189885929226875
step: 110, loss: 0.0014459651429206133
step: 120, loss: 0.1699659824371338
step: 130, loss: 0.005954301916062832
step: 140, loss: 0.004011765122413635
step: 150, loss: 0.0029189579654484987
step: 160, loss: 0.00485085416585207
step: 170, loss: 0.025249846279621124
step: 180, loss: 0.07731613516807556
step: 190, loss: 0.028590409085154533
step: 200, loss: 0.0004531518497969955
step: 210, loss: 0.0035903374664485455
```

```

step: 220, loss: 0.005076893605291843
step: 230, loss: 0.004562031477689743
step: 240, loss: 0.03050367906689644
step: 250, loss: 0.06842093169689178
step: 260, loss: 0.011338386684656143
step: 270, loss: 0.012906242161989212
step: 280, loss: 0.0067929429933428764
step: 290, loss: 0.002001001965254545
step: 300, loss: 0.0006495245033875108
step: 310, loss: 0.015843907371163368
step: 320, loss: 0.1213579773902893
step: 330, loss: 0.06239492818713188
step: 340, loss: 0.05767865106463432
step: 350, loss: 0.0025573582388460636
step: 360, loss: 0.02961813285946846
step: 370, loss: 0.08889074623584747
step: 380, loss: 0.017654744908213615
step: 390, loss: 0.002834367100149393
step: 400, loss: 0.061882201582193375
step: 410, loss: 0.002868249546736479
step: 420, loss: 0.011381510645151138
step: 430, loss: 0.004690353758633137
step: 440, loss: 0.05785266309976578
step: 450, loss: 0.02484436146914959
step: 460, loss: 0.043180085718631744
step: 470, loss: 0.1760336309671402
step: 480, loss: 0.003785078413784504
step: 490, loss: 0.006956083234399557
step: 500, loss: 0.0025965292006731033
step: 510, loss: 0.000906239845789969
step: 520, loss: 0.00246328953653574
step: 530, loss: 0.006275826599448919
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9328426862925483, f1=0.9238625812441968,
best_f1=0.9238625812441968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003383631119504571
step: 10, loss: 0.002862591529265046
step: 20, loss: 0.07956745475530624
step: 30, loss: 0.010443693958222866

```

step: 40, loss: 0.0008485241560265422
step: 50, loss: 0.02073807828128338
step: 60, loss: 0.0012124389177188277
step: 70, loss: 0.0032604001462459564
step: 80, loss: 0.007614240515977144
step: 90, loss: 0.0041191186755895615
step: 100, loss: 0.0019345306791365147
step: 110, loss: 0.0006252852035686374
step: 120, loss: 0.0011670986423268914
step: 130, loss: 0.0041665309108793736
step: 140, loss: 0.000540204462595284
step: 150, loss: 0.0008390029543079436
step: 160, loss: 0.0028693799395114183
step: 170, loss: 0.06681474298238754
step: 180, loss: 0.0006178351468406618
step: 190, loss: 0.005187974311411381
step: 200, loss: 0.0025763693265616894
step: 210, loss: 0.05857887491583824
step: 220, loss: 0.0016154606128111482
step: 230, loss: 0.20945870876312256
step: 240, loss: 0.06972826272249222
step: 250, loss: 0.0016729983035475016
step: 260, loss: 0.0022899569012224674
step: 270, loss: 0.024076661095023155
step: 280, loss: 0.007038275711238384
step: 290, loss: 0.0017550303600728512
step: 300, loss: 8.636529673822224e-05
step: 310, loss: 0.0007058375631459057
step: 320, loss: 0.0009053483372554183
step: 330, loss: 0.0012769645545631647
step: 340, loss: 0.027121735736727715
step: 350, loss: 0.00017139024566859007
step: 360, loss: 0.21972550451755524
step: 370, loss: 0.1857607066631317
step: 380, loss: 0.0004243631847202778
step: 390, loss: 0.011023646220564842
step: 400, loss: 0.0008448585285805166
step: 410, loss: 0.008730070665478706
step: 420, loss: 0.0002970362256746739
step: 430, loss: 0.0038857413455843925
step: 440, loss: 0.024444419890642166
step: 450, loss: 0.0020603027660399675
step: 460, loss: 0.0026204739697277546
step: 470, loss: 0.05342244729399681
step: 480, loss: 0.08212205767631531
step: 490, loss: 0.03532245755195618
step: 500, loss: 0.014811329543590546
step: 510, loss: 0.0039772531017661095

```

step: 520, loss: 0.00041172452620230615
step: 530, loss: 0.00134041637647897
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9360902255639098, f1=0.9219924812030075,
best_f1=0.9219924812030075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007550226873718202
step: 10, loss: 0.07658907026052475
step: 20, loss: 0.014071359299123287
step: 30, loss: 0.06836313009262085
step: 40, loss: 0.003712432226166129
step: 50, loss: 0.0029671103693544865
step: 60, loss: 0.0011604587780311704
step: 70, loss: 0.0008313973667100072
step: 80, loss: 0.008306872099637985
step: 90, loss: 0.040910009294748306
step: 100, loss: 0.0005221553146839142
step: 110, loss: 0.02820379100739956
step: 120, loss: 0.18094663321971893
step: 130, loss: 0.0037822872400283813
step: 140, loss: 0.014697112143039703
step: 150, loss: 0.007604703307151794
step: 160, loss: 0.004449903499335051
step: 170, loss: 0.012736300937831402
step: 180, loss: 0.006604826543480158
step: 190, loss: 0.1470205932855606
step: 200, loss: 0.008633404970169067
step: 210, loss: 0.005755546968430281
step: 220, loss: 0.0014522764831781387
step: 230, loss: 0.0007187661831267178
step: 240, loss: 0.009669559076428413
step: 250, loss: 0.08055216073989868
step: 260, loss: 0.005136456806212664
step: 270, loss: 0.012073181569576263
step: 280, loss: 0.0009799712570384145
step: 290, loss: 0.0033185824286192656
step: 300, loss: 0.00014212164387572557
step: 310, loss: 0.06912513822317123
step: 320, loss: 0.0008994760573841631
step: 330, loss: 0.03186517953872681

```

```

step: 340, loss: 0.07904764264822006
step: 350, loss: 0.025856604799628258
step: 360, loss: 0.00385713460855186
step: 370, loss: 0.004922663327306509
step: 380, loss: 0.00246004993095994
step: 390, loss: 0.0003457654966041446
step: 400, loss: 0.09330029785633087
step: 410, loss: 0.0022417400032281876
step: 420, loss: 0.0091751329600811
step: 430, loss: 0.023916171863675117
step: 440, loss: 6.61923477309756e-05
step: 450, loss: 0.01091964915394783
step: 460, loss: 0.0004689162306021899
step: 470, loss: 0.000756095047108829
step: 480, loss: 9.650007268646732e-05
step: 490, loss: 0.002571765100583434
step: 500, loss: 0.022696755826473236
step: 510, loss: 0.015153964050114155
step: 520, loss: 0.020009523257613182
step: 530, loss: 0.006845493800938129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9398148148148149, f1=0.9300797747536368,
best_f1=0.9300797747536368
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07712079584598541
step: 10, loss: 0.0009342145640403032
step: 20, loss: 0.0073060099966824055
step: 30, loss: 0.002729887142777443
step: 40, loss: 0.021951425820589066
step: 50, loss: 0.0012916064588353038
step: 60, loss: 0.003644883166998625
step: 70, loss: 0.0014888579025864601
step: 80, loss: 0.1374431848526001
step: 90, loss: 0.008188922889530659
step: 100, loss: 0.020244134590029716
step: 110, loss: 0.019766151905059814
step: 120, loss: 0.001437539467588067
step: 130, loss: 0.00026116776280105114
step: 140, loss: 0.0008014136692509055
step: 150, loss: 0.001133420504629612

```

```

step: 160, loss: 0.0002842065005097538
step: 170, loss: 0.004542922601103783
step: 180, loss: 0.002230622572824359
step: 190, loss: 0.0012900776928290725
step: 200, loss: 0.0012179811019450426
step: 210, loss: 0.006347783841192722
step: 220, loss: 0.0009653966990299523
step: 230, loss: 0.0003598715993575752
step: 240, loss: 0.007757126819342375
step: 250, loss: 0.021244650706648827
step: 260, loss: 0.01726342737674713
step: 270, loss: 0.0012922721216455102
step: 280, loss: 0.026976022869348526
step: 290, loss: 0.003421657020226121
step: 300, loss: 0.00027842819690704346
step: 310, loss: 0.003756972262635827
step: 320, loss: 0.002509981393814087
step: 330, loss: 0.12650847434997559
step: 340, loss: 0.004255066625773907
step: 350, loss: 0.001101645058952272
step: 360, loss: 0.0003175330930389464
step: 370, loss: 0.0023567217867821455
step: 380, loss: 0.0072877597995102406
step: 390, loss: 0.000253694859566167
step: 400, loss: 0.0035071424208581448
step: 410, loss: 0.0008663029875606298
step: 420, loss: 0.0022020465694367886
step: 430, loss: 0.0020804598461836576
step: 440, loss: 0.00033287302358075976
step: 450, loss: 0.05120720714330673
step: 460, loss: 0.0026103465352207422
step: 470, loss: 0.022233974188566208
step: 480, loss: 0.24847544729709625
step: 490, loss: 0.0038210095372051
step: 500, loss: 0.0024386760778725147
step: 510, loss: 0.00018704916874412447
step: 520, loss: 0.004939771723002195
step: 530, loss: 0.028408266603946686
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9429622815087396, f1=0.9322191272051997,
best_f1=0.9322191272051997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.00013630383182317019
step: 10, loss: 0.0023037134669721127
step: 20, loss: 0.00022074556909501553
step: 30, loss: 0.00019101832003798336
step: 40, loss: 0.0005326076061464846
step: 50, loss: 0.0004674845840781927
step: 60, loss: 0.0017545666778460145
step: 70, loss: 0.0038672671653330326
step: 80, loss: 0.004872037563472986
step: 90, loss: 0.025887761265039444
step: 100, loss: 0.0035187348257750273
step: 110, loss: 0.0005869740853086114
step: 120, loss: 0.0003758080711122602
step: 130, loss: 0.0006755513604730368
step: 140, loss: 0.0014130283379927278
step: 150, loss: 0.0014217214193195105
step: 160, loss: 0.000405714672524482
step: 170, loss: 0.0002885662252083421
step: 180, loss: 0.00011517976236063987
step: 190, loss: 0.0009772235061973333
step: 200, loss: 0.0018752175383269787
step: 210, loss: 0.0018042929004877806
step: 220, loss: 0.10132656246423721
step: 230, loss: 0.0006025477196089923
step: 240, loss: 0.01793835498392582
step: 250, loss: 0.0014097130624577403
step: 260, loss: 0.0028914534486830235
step: 270, loss: 0.022234657779335976
step: 280, loss: 0.0018388121388852596
step: 290, loss: 0.005365130957216024
step: 300, loss: 0.009822693653404713
step: 310, loss: 0.003006704617291689
step: 320, loss: 0.003028400707989931
step: 330, loss: 3.473741526249796e-05
step: 340, loss: 0.010272473096847534
step: 350, loss: 0.0011238326551392674
step: 360, loss: 0.0012848349288105965
step: 370, loss: 0.001184965600259602
step: 380, loss: 0.009496178478002548
step: 390, loss: 0.0033636484295129776
step: 400, loss: 0.00034707196755334735
step: 410, loss: 0.0009055056143552065
step: 420, loss: 0.03075364977121353
step: 430, loss: 0.002573159523308277
step: 440, loss: 0.0008040462271310389
step: 450, loss: 0.0012767196167260408
```



```

step: 460, loss: 0.006141608115285635
step: 470, loss: 0.00048537817201577127
step: 480, loss: 0.003099660621955991
step: 490, loss: 0.0003140986373182386
step: 500, loss: 0.02946791984140873
step: 510, loss: 0.0019926836248487234
step: 520, loss: 0.007211483549326658
step: 530, loss: 0.011199313215911388
/usr/local/lib/python3.7/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=0.9342789598108747, f1=0.9168241965973536,
best_f1=0.9322191272051997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00533313350751996
step: 10, loss: 0.008364317938685417
step: 20, loss: 0.016890423372387886
step: 30, loss: 5.8961304603144526e-05
step: 40, loss: 0.04072338342666626
step: 50, loss: 0.001750967581756413
step: 60, loss: 7.409586396533996e-05
step: 70, loss: 0.0066320206969976425
step: 80, loss: 0.00033461753628216684
step: 90, loss: 0.0009770009201020002
step: 100, loss: 0.0025217835791409016
step: 110, loss: 0.017548084259033203
step: 120, loss: 0.06966578215360641
step: 130, loss: 0.0009508402436040342
step: 140, loss: 0.00014307697711046785
step: 150, loss: 0.0024089571088552475
step: 160, loss: 0.0067294687032699585
step: 170, loss: 0.0003442400775384158
step: 180, loss: 0.0001254429662367329
step: 190, loss: 0.0005174994003027678
step: 200, loss: 0.0017778398469090462
step: 210, loss: 0.00020452849275898188
step: 220, loss: 0.0007964679971337318
step: 230, loss: 0.0009020571014843881
step: 240, loss: 0.0012885159812867641
step: 250, loss: 4.492099105846137e-05
step: 260, loss: 0.018030572682619095
step: 270, loss: 0.001343769719824195
step: 280, loss: 0.001493748277425766
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 290, loss: 0.002471392508596182
step: 300, loss: 0.0006738342926837504

```

```

step: 310, loss: 0.0011102674761787057
step: 320, loss: 0.1373155564069748
step: 330, loss: 0.0013428680831566453
step: 340, loss: 0.010961201973259449
step: 350, loss: 0.00018872454529628158
step: 360, loss: 0.0004029158444609493
step: 370, loss: 0.0002550210920162499
step: 380, loss: 0.00012804627476725727
step: 390, loss: 0.0009064827463589609
step: 400, loss: 0.0005835273186676204
step: 410, loss: 0.0006464730831794441
step: 420, loss: 0.00153865956235677
step: 430, loss: 0.088473379611969
step: 440, loss: 0.039003003388643265
step: 450, loss: 0.048637934029102325
step: 460, loss: 0.0005073624779470265
step: 470, loss: 0.0004874256846960634
step: 480, loss: 0.05564360320568085
step: 490, loss: 0.0021508694626390934
step: 500, loss: 5.019831223762594e-05
step: 510, loss: 0.15473227202892303
step: 520, loss: 0.0029149448964744806
step: 530, loss: 0.009454556740820408
/usr/local/lib/python3.7/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=0.93790546802595, f1=0.9150812064965197,
best_f1=0.9322191272051997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002845074748620391
step: 10, loss: 0.0021921403240412474
step: 20, loss: 0.004446128848940134
step: 30, loss: 0.00026173132937401533
step: 40, loss: 0.056037940084934235
step: 50, loss: 0.0012653105659410357
step: 60, loss: 0.0023466411512345076
step: 70, loss: 0.0007194370264187455
step: 80, loss: 0.001290071988478303
step: 90, loss: 0.00026398716727271676
step: 100, loss: 0.002941135549917817
step: 110, loss: 0.004785514902323484
step: 120, loss: 0.0005029794410802424
step: 130, loss: 0.006274931598454714
step: 140, loss: 0.0038992278277873993
step: 150, loss: 0.003682374954223633
step: 160, loss: 0.0007076628389768302

```

```

step: 170, loss: 0.00035214071976952255
step: 180, loss: 0.0023112159688025713
step: 190, loss: 0.0006747624720446765
step: 200, loss: 0.0001726112823234871
step: 210, loss: 0.0004285327740944922
step: 220, loss: 0.02533683180809021
step: 230, loss: 0.025034157559275627
step: 240, loss: 0.005222815088927746
step: 250, loss: 0.007708042860031128
step: 260, loss: 0.05289643257856369
step: 270, loss: 0.019407758489251137
step: 280, loss: 0.01627064123749733
step: 290, loss: 0.010652377270162106
step: 300, loss: 0.0013239027466624975
step: 310, loss: 0.02971455827355385
step: 320, loss: 0.0009159466717392206
step: 330, loss: 0.007060125470161438
step: 340, loss: 0.0019808150827884674
step: 350, loss: 0.00019739984418265522
step: 360, loss: 0.13448548316955566
step: 370, loss: 0.0037692987825721502
step: 380, loss: 0.0003642031515482813
step: 390, loss: 0.008544796146452427
step: 400, loss: 6.65643165120855e-05
step: 410, loss: 0.001591612584888935
step: 420, loss: 0.00032516499049961567
step: 430, loss: 0.0017133635701611638
step: 440, loss: 0.00027398165548220277
step: 450, loss: 0.0005686373915523291
step: 460, loss: 0.015769343823194504
step: 470, loss: 0.057448141276836395
step: 480, loss: 7.660088158445433e-05
step: 490, loss: 7.938459020806476e-05
step: 500, loss: 0.0002716774761211127
step: 510, loss: 0.00040006908238865435
step: 520, loss: 0.0008442626567557454
step: 530, loss: 0.00017119647236540914
/usr/local/lib/python3.7/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=0.9381918819188192, f1=0.9238625812441968,
best_f1=0.9322191272051997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009879304561764002
step: 10, loss: 0.0409211628139019
step: 20, loss: 0.0003243899845983833

```

step: 30, loss: 0.002953212708234787
step: 40, loss: 0.0007990446756593883
step: 50, loss: 0.0018270846921950579
step: 60, loss: 0.0007527304696850479
step: 70, loss: 0.001102553796954453
step: 80, loss: 0.0008122518192976713
step: 90, loss: 2.4228465917985886e-05
step: 100, loss: 0.0006047065835446119
step: 110, loss: 0.00119203538633883
step: 120, loss: 8.659171726321802e-05
step: 130, loss: 0.00048155777039937675
step: 140, loss: 0.00647902162745595
step: 150, loss: 0.08935816586017609
step: 160, loss: 0.00043896635179407895
step: 170, loss: 0.001586464000865817
step: 180, loss: 0.0002822738897521049
step: 190, loss: 0.0005628187791444361
step: 200, loss: 0.00017005595145747066
step: 210, loss: 0.00026457649073563516
step: 220, loss: 0.006959948223084211
step: 230, loss: 0.013441173359751701
step: 240, loss: 0.0005915469373576343
step: 250, loss: 0.05626928433775902
step: 260, loss: 0.0006088150548748672
step: 270, loss: 0.0011654106201604009
step: 280, loss: 0.0030765000265091658
step: 290, loss: 0.0001300408475799486
step: 300, loss: 0.0005098533583804965
step: 310, loss: 0.00045507538015954196
step: 320, loss: 9.244819375453517e-05
step: 330, loss: 0.0023148218169808388
step: 340, loss: 0.0002589290670584887
step: 350, loss: 0.0001190434632007964
step: 360, loss: 0.000567737384699285
step: 370, loss: 0.002546025672927499
step: 380, loss: 0.05360597372055054
step: 390, loss: 0.0024286473635584116
step: 400, loss: 0.000889384129550308
step: 410, loss: 5.4397394706029445e-05
step: 420, loss: 0.0002550631470512599
step: 430, loss: 0.0009760463144630194
step: 440, loss: 0.0030378533992916346
step: 450, loss: 0.00011174689279869199
step: 460, loss: 0.0014172218507155776
step: 470, loss: 2.7382640837458894e-05
step: 480, loss: 0.0002950281777884811
step: 490, loss: 0.00014982644643168896
step: 500, loss: 0.0016838170122355223

```

step: 510, loss: 0.0006114077405072749
step: 520, loss: 0.00040078675374388695
step: 530, loss: 0.0042498186230659485
/usr/local/lib/python3.7/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=0.936768149882904, f1=0.9220657276995305,
best_f1=0.9322191272051997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016632068436592817
step: 10, loss: 0.004829268902540207
step: 20, loss: 0.00208502565510571
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 30, loss: 0.0018520207377150655
step: 40, loss: 9.403594594914466e-05
step: 50, loss: 0.01594553142786026
step: 60, loss: 0.0007238123216666281
step: 70, loss: 0.00023893029720056802
step: 80, loss: 0.00017591517826076597
step: 90, loss: 0.004168366547673941
step: 100, loss: 0.00035840016789734364
step: 110, loss: 0.1560969054698944
step: 120, loss: 0.00015018830890767276
step: 130, loss: 0.0009123368072323501
step: 140, loss: 0.00011208041541976854
step: 150, loss: 0.0001473795564379543
step: 160, loss: 0.0007326003978960216
step: 170, loss: 0.00016919424524530768
step: 180, loss: 0.001167413662187755
step: 190, loss: 0.15447324514389038
step: 200, loss: 0.004593167454004288
step: 210, loss: 0.0004836739390157163
step: 220, loss: 0.0007663048454560339
step: 230, loss: 0.0006338538369163871
step: 240, loss: 7.419844041578472e-05
step: 250, loss: 0.00010085565008921549
step: 260, loss: 0.0007516191690228879
step: 270, loss: 0.0001318736613029614
step: 280, loss: 0.00023405755928251892
step: 290, loss: 0.0004260671266820282
step: 300, loss: 0.00024878286058083177
step: 310, loss: 0.001648034667596221
step: 320, loss: 5.8909419749397784e-05
step: 330, loss: 0.0001733538811095059
step: 340, loss: 0.0001818334130803123
step: 350, loss: 0.002426623133942485

```

```

step: 360, loss: 0.002203878480941057
step: 370, loss: 0.0016672926722094417
step: 380, loss: 0.0004676340613514185
step: 390, loss: 8.984347368823364e-05
step: 400, loss: 0.0031588857527822256
step: 410, loss: 0.00031392942764796317
step: 420, loss: 0.0010376228019595146
step: 430, loss: 0.0002815687912516296
step: 440, loss: 0.0009443257004022598
step: 450, loss: 0.0010048376861959696
step: 460, loss: 0.0010841755429282784
step: 470, loss: 0.0005747207906097174
step: 480, loss: 0.002274223603308201
step: 490, loss: 0.0005162042798474431
step: 500, loss: 0.0008546155295334756
step: 510, loss: 5.557648910325952e-05
step: 520, loss: 3.2511856261407956e-05
step: 530, loss: 0.008819792419672012
/usr/local/lib/python3.7/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=0.9372967951695309, f1=0.922071861875875,
best_f1=0.9322191272051997
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```
An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-  
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
5742it [00:35, 163.39it/s]  
load_f1 = 0.9417206290471785  
real_f1 = 0.9404706968158745  
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-  
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.  
  warnings.warn("An input tensor was not cuda.")  
4398it [00:30, 144.95it/s]
```