

FMedium_50_1_2_bert

August 13, 2022

0.1 Install packages

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

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

```
Cloning into 'ditto'...
remote: Enumerating objects: 291, done.
remote: Total 291 (delta 0), reused 0 (delta 0), pack-reused 291
Receiving objects: 100% (291/291), 26.87 MiB | 8.64 MiB/s, done.
Resolving deltas: 100% (142/142), done.
Checking out files: 100% (128/128), done.
/content/ditto
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Downloading gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
    |                               | 24.2 MB 1.4 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 13.3 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 65.5 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 57.7 MB/s
Requirement already satisfied: sklearn==0.0 in
```

```

/usr/local/lib/python3.7/dist-packages (from -r requirements.txt (line 6)) (0.0)
Collecting spacy==3.1
  Downloading
spacy-3.1.0-cp37-cp37m-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (6.4 MB)
  | 6.4 MB 23.9 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 | 8.47 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 3.0 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 67.5 MB/s
Collecting torch==1.4.0
  Downloading torch-1.4.0-cp37-cp37m-manylinux1_x86_64.whl (753.4 MB)
  | 753.4 MB 6.9 kB/s
Collecting tqdm==4.41.0
  Downloading tqdm-4.41.0-py2.py3-none-any.whl (56 kB)
  | 56 kB 5.0 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
  | 884 kB 38.1 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 48.0 MB/s
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: six>=1.5.0 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.15.0)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
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: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Requirement already satisfied: 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 plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
Collecting thinc<7.4.0,>=7.3.0
    Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
      |                               | 2.2 MB 55.6 MB/s
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                               | 184 kB 75.3 MB/s
Collecting blis<0.5.0,>=0.4.0
    Downloading blis-0.4.1-cp37-cp37m-manylinux1_x86_64.whl (3.7 MB)
      |                               | 3.7 MB 26.3 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: 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: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.23.0)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting tokenizers==0.8.1.rc2
    Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
      |                               | 3.0 MB 52.4 MB/s

```

```

Requirement already satisfied: packaging in /usr/local/lib/python3.7/dist-
packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Collecting sacremoses
  Downloading sacremoses-0.0.53.tar.gz (880 kB)
    |                                     | 880 kB 58.6 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: 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: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line
4)) (2.10)
Requirement already satisfied: 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=1449926
    sha256=9a4ec656142c26492480bd0d2598d97ae9fd8e16291a92f66fd6e88bf9463ab5
    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=cdab4d41c25e32758ec7f054c6977dbc9a49debdfe91257cc0435a3a99610245
  Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
  Attempting uninstall: numpy
    Found existing installation: numpy 1.21.6
    Uninstalling numpy-1.21.6:
      Successfully uninstalled numpy-1.21.6
  Attempting uninstall: tqdm
    Found existing installation: tqdm 4.64.0
    Uninstalling tqdm-4.64.0:
      Successfully uninstalled tqdm-4.64.0
  Attempting uninstall: srsly
    Found existing installation: srsly 2.4.4
    Uninstalling srsly-2.4.4:
      Successfully uninstalled srsly-2.4.4
  Attempting uninstall: regex
    Found existing installation: regex 2022.6.2
    Uninstalling regex-2022.6.2:
      Successfully uninstalled regex-2022.6.2
  Attempting uninstall: blis
    Found existing installation: blis 0.7.8
    Uninstalling blis-0.7.8:
      Successfully uninstalled blis-0.7.8
  Attempting uninstall: thinc
    Found existing installation: thinc 8.1.0
    Uninstalling thinc-8.1.0:
      Successfully uninstalled thinc-8.1.0
  Attempting uninstall: catalogue
    Found existing installation: catalogue 2.0.8
    Uninstalling catalogue-2.0.8:
      Successfully uninstalled catalogue-2.0.8
  Attempting uninstall: torch
    Found existing installation: torch 1.12.0+cu113
    Uninstalling torch-1.12.0+cu113:
      Successfully uninstalled torch-1.12.0+cu113
  Attempting uninstall: spacy
    Found existing installation: spacy 3.4.1
    Uninstalling spacy-3.4.1:
      Successfully uninstalled spacy-3.4.1
  Attempting uninstall: nltk
    Found existing installation: nltk 3.7
    Uninstalling nltk-3.7:
      Successfully uninstalled nltk-3.7
```

```

Attempting uninstall: gensim
Found existing installation: gensim 3.6.0
Uninstalling gensim-3.6.0:
  Successfully uninstalled gensim-3.6.0
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.
xarray-einstats 0.2.2 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
torchvision 0.13.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.
torchtext 0.13.0 requires torch==1.12.0, but you have torch 1.4.0 which is
incompatible.
torchaudio 0.12.0+cu113 requires torch==1.12.0, but you have torch 1.4.0 which
is incompatible.
tensorflow 2.8.2+zzzcolab20220719082949 requires numpy>=1.20, but you have numpy
1.19.2 which is incompatible.
panel 0.12.1 requires tqdm>=4.48.0, but you have tqdm 4.41.0 which is
incompatible.
fastai 2.7.7 requires torch<1.13,>=1.7, but you have torch 1.4.0 which is
incompatible.
en-core-web-sm 3.4.0 requires spacy<3.5.0,>=3.4.0, but you have spacy 2.2.3
which is incompatible.
cmdstanpy 1.0.4 requires numpy>=1.21, but you have numpy 1.19.2 which is
incompatible.
Successfully installed blis-0.4.1 catalogue-1.0.0 gensim-3.8.1 jsonlines-1.2.0
nltk-3.4.5 numpy-1.19.2 plac-1.1.3 regex-2019.12.20 sacremoses-0.0.53
sentencepiece-0.1.85 spacy-2.2.3 srsly-1.0.5 tensorboardX-2.0 thinc-7.3.1
tokenizers-0.8.1rc2 torch-1.4.0 tqdm-4.41.0 transformers-3.1.0
/content/ditto
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Unzipping corpora/stopwords.zip.

[1]: True

```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10117, done.
remote: Counting objects: 100% (233/233), done.
remote: Compressing objects: 100% (165/165), done.
remote: Total 10117 (delta 118), reused 152 (delta 62), pack-reused 9884
Receiving objects: 100% (10117/10117), 14.96 MiB | 12.44 MiB/s, done.
Resolving deltas: 100% (6922/6922), 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-7p22dgpv
Created temporary directory: /tmp/pip-req-tracker-7_oqwpxi
Initialized build tracking at /tmp/pip-req-tracker-7_oqwpxi
Created build tracker: /tmp/pip-req-tracker-7_oqwpxi
Entered build tracker: /tmp/pip-req-tracker-7_oqwpxi
```

```

Created temporary directory: /tmp/pip-install-16x4nzkv
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-e3egxi5z
  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-7_oqwpxi'
  Running setup.py (path:/tmp/pip-req-build-e3egxi5z/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-l29pz39a
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-l29pz39a/apex.egg-info
  writing /tmp/pip-pip-egg-info-l29pz39a/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-l29pz39a/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-l29pz39a/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-l29pz39a/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-l29pz39a/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-e3egxi5z 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-7_oqwpxi'
Created temporary directory: /tmp/pip-unpack-2nkwys05
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-da5dr70s
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-da5dr70s
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = """/tmp/pip-req-build-e3egxi5z/setup.py""';

```



```
__file__='''/tmp/pip-req-build-e3egxi5z/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-da5dr70s
```

```
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/parallel
copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
copying apex/parallel/sync_batchnorm_kernel.py -> build/lib/apex/parallel
copying apex/parallel/optimized_sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/multiproc.py -> build/lib/apex/parallel
copying apex/parallel/distributed.py -> build/lib/apex/parallel
copying apex/parallel/__init__.py -> build/lib/apex/parallel
copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
copying apex/parallel/LARC.py -> build/lib/apex/parallel
creating build/lib/apex/RNN
copying apex/RNN/models.py -> build/lib/apex/RNN
copying apex/RNN/cells.py -> build/lib/apex/RNN
copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
copying apex/RNN/__init__.py -> build/lib/apex/RNN
creating build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/__init__.py ->
build/lib/apex/multi_tensor_apply
copying apex/multi_tensor_apply/multi_tensor_apply.py ->
build/lib/apex/multi_tensor_apply
creating build/lib/apex/transformer
copying apex/transformer/enums.py -> build/lib/apex/transformer
copying apex/transformer/log_util.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/__init__.py -> build/lib/apex/transformer
copying apex/transformer/parallel_state.py -> build/lib/apex/transformer
creating build/lib/apex/optimizers
copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
```

```

copying apex/optimizers/fused_sgd.py -> build/lib/apex/optimizers
copying apex/optimizers/__init__.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_adam.py -> build/lib/apex/optimizers
copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
creating build/lib/apex/fp16_utils
copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
creating build/lib/apex/contrib
copying apex/contrib/__init__.py -> build/lib/apex/contrib
creating build/lib/apex/amp
copying apex/amp/wrap.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
copying apex/amp/_initialize.py -> build/lib/apex/amp
copying apex/amp/_amp_state.py -> build/lib/apex/amp
copying apex/amp/opt.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/__init__.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/scaler.py -> build/lib/apex/amp
creating build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
copying apex/normalization/__init__.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
creating build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/_timers.py ->
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/utils.py ->
build/lib/apex/transformer/pipeline_parallel
copying apex/transformer/pipeline_parallel/__init__.py ->
build/lib/apex/transformer/pipeline_parallel
creating build/lib/apex/transformer/layers
copying apex/transformer/layers/layer_norm.py ->

```

```

build/lib/apex/transformer/layers
  copying apex/transformer/layers/__init__.py ->
build/lib/apex/transformer/layers
  creating build/lib/apex/transformer/_data
  copying apex/transformer/_data/_batch_sampler.py ->
build/lib/apex/transformer/_data
  copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
  creating 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/global_vars.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
  copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing
  creating build/lib/apex/transformer/functional
  copying apex/transformer/functional/__init__.py ->
build/lib/apex/transformer/functional
  copying apex/transformer/functional/fused_softmax.py ->
build/lib/apex/transformer/functional
  creating build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/mappings.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/data.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/cross_entropy.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
  copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
  creating build/lib/apex/transformer/amp
  copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp

```

```

    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    creating 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_no_pipelining.py
-> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/__init__.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    copying apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py -> build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.py -> build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    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/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/layer_norm.py ->
build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/__init__.py ->
build/lib/apex/contrib/layer_norm
    creating build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/fast_encdec_multihead_attn_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/fast_self_multihead_attn_norm_add_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/__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/fast_self_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
    creating build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    creating 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/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
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/fused_adam.py ->
build/lib/apex/contrib/optimizers
    creating build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/__init__.py ->
build/lib/apex/contrib/transducer
    copying apex/contrib/transducer/transducer.py ->

```

```

build/lib/apex/contrib/transducer
  creating build/lib/apex/contrib/index_mul_2d
  copying apex/contrib/index_mul_2d/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/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchanger_1d.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/__init__.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_halo_exchange_module_tests.py ->
build/lib/apex/contrib/peer_memory
  copying apex/contrib/peer_memory/peer_memory.py ->
build/lib/apex/contrib/peer_memory
  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/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
  copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
  creating build/lib/apex/amp/lists
  copying apex/amp/lists/functional_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
  copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
  installing to build/bdist.linux-x86_64/wheel
  running install
  running install_lib
  creating build/bdist.linux-x86_64
  creating build/bdist.linux-x86_64/wheel
  creating build/bdist.linux-x86_64/wheel/apex
  creating build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/optimized_sync_batchnorm.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
  copying build/lib/apex/parallel/multiproc.py ->
build/bdist.linux-x86_64/wheel/apex/parallel

```

```

    copying build/lib/apex/parallel/distributed.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.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
    copying build/lib/apex/parallel/LARC.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    creating 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/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    copying build/lib/apex/multi_tensor_apply/multi_tensor_apply.py ->
build/bdist.linux-x86_64/wheel/apex/multi_tensor_apply
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/enums.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/pipeline_parallel
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/schedules/common.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/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/__init__.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/p2p_communication.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/pipeline_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating build/bdist.linux-x86_64/wheel/apex/transformer/layers
    copying build/lib/apex/transformer/layers/layer_norm.py ->
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
    creating 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
    copying build/lib/apex/transformer/_data/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    creating 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/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_bert.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/distributed_test_base.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/microbatches.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/functional/fused_softmax.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/functional
    copying build/lib/apex/transformer/utils.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/mappings.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/data.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/cross_entropy.py ->

```



```

build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/random.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
  copying build/lib/apex/transformer/tensor_parallel/utils.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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
  copying build/lib/apex/transformer/amp/grad_scaler.py ->
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/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
  creating build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_sgd.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_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  copying build/lib/apex/optimizers/fused_mixed_precision_lamb.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
  creating build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  copying build/lib/apex/fp16_utils/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
  creating build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  creating 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/__init__.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
  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/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  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/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
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
  creating 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
  copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
  creating 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
  copying build/lib/apex/contrib/layer_norm/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/layer_norm
  creating build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_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/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/encdec_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/fast_self_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
    creating 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
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating 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/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/permutation_lib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating 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/distributed_fused_lamb.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/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/fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers

```

```

    creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
    copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
    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/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/__init__.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
    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/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    creating 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/__version__.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/compat.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/_initialize.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/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/_process_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/utils.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/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
    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/__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/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating 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/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
    creating build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/mlp
    copying build/lib/apex/mlp/mlp.py -> build/bdist.linux-x86_64/wheel/apex/mlp
    creating 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
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    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-da5dr70s/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
    adding 'apex/__init__.py'
    adding 'apex/_autocast_utils.py'
    adding 'apex/RNN/RNNBackend.py'

```

```

adding 'apex/RNN/__init__.py'
adding 'apex/RNN/cells.py'
adding 'apex/RNN/models.py'
adding 'apex/amp/__init__.py'
adding 'apex/amp/__version__.py'
adding 'apex/amp/_amp_state.py'
adding 'apex/amp/_initialize.py'
adding 'apex/amp/_process_optimizer.py'
adding 'apex/amp/amp.py'
adding 'apex/amp/compat.py'
adding 'apex/amp/frontend.py'
adding 'apex/amp/handle.py'
adding 'apex/amp/opt.py'
adding 'apex/amp/rnn_compat.py'
adding 'apex/amp/scaler.py'
adding 'apex/amp/utils.py'
adding 'apex/amp/wrap.py'
adding 'apex/amp/lists/__init__.py'
adding 'apex/amp/lists/functional_overrides.py'
adding 'apex/amp/lists/tensor_overrides.py'
adding 'apex/amp/lists/torch_overrides.py'
adding 'apex/contrib/__init__.py'
adding 'apex/contrib/bottleneck/__init__.py'
adding 'apex/contrib/bottleneck/bottleneck.py'
adding 'apex/contrib/bottleneck/bottleneck_module_test.py'
adding 'apex/contrib/bottleneck/halo_exchangers.py'
adding 'apex/contrib/bottleneck/test.py'
adding 'apex/contrib/clip_grad/__init__.py'
adding 'apex/contrib/clip_grad/clip_grad.py'
adding 'apex/contrib/conv_bias_relu/__init__.py'
adding 'apex/contrib/conv_bias_relu/conv_bias_relu.py'
adding 'apex/contrib/fmha/__init__.py'
adding 'apex/contrib/fmha/fmha.py'
adding 'apex/contrib/focal_loss/__init__.py'
adding 'apex/contrib/focal_loss/focal_loss.py'
adding 'apex/contrib/groupbn/__init__.py'
adding 'apex/contrib/groupbn/batch_norm.py'
adding 'apex/contrib/index_mul_2d/__init__.py'
adding 'apex/contrib/index_mul_2d/index_mul_2d.py'
adding 'apex/contrib/layer_norm/__init__.py'
adding 'apex/contrib/layer_norm/layer_norm.py'
adding 'apex/contrib/multihead_attn/__init__.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn.py'
adding 'apex/contrib/multihead_attn/encdec_multihead_attn_func.py'
adding 'apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py'
adding
'apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_add_func.py'
    adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_func.py'

```

```

adding 'apex/contrib/multihead_attn/fast_self_multihead_attn_norm_add_func.py'
adding 'apex/contrib/multihead_attn/mask_softmax_dropout_func.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn.py'
adding 'apex/contrib/multihead_attn/self_multihead_attn_func.py'
adding 'apex/contrib/optimizers/__init__.py'
adding 'apex/contrib/optimizers/distributed_fused_adam.py'
adding 'apex/contrib/optimizers/distributed_fused_lamb.py'
adding 'apex/contrib/optimizers/fp16_optimizer.py'
adding 'apex/contrib/optimizers/fused_adam.py'
adding 'apex/contrib/optimizers/fused_lamb.py'
adding 'apex/contrib/optimizers/fused_sgd.py'
adding 'apex/contrib/peer_memory/__init__.py'
adding 'apex/contrib/peer_memory/peer_halo_exchange_module_tests.py'
adding 'apex/contrib/peer_memory/peer_halo_exchanger_1d.py'
adding 'apex/contrib/peer_memory/peer_memory.py'
adding 'apex/contrib/sparsity/__init__.py'
adding 'apex/contrib/sparsity/asp.py'
adding 'apex/contrib/sparsity/permutation_lib.py'
adding 'apex/contrib/sparsity/sparse_masklib.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/__init__.py'
adding 'apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py'
    adding 'apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py'
    adding
'apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py'
    adding 'apex/contrib/transducer/__init__.py'
    adding 'apex/contrib/transducer/transducer.py'
    adding 'apex/contrib/xentropy/__init__.py'
    adding 'apex/contrib/xentropy/softmax_xentropy.py'
    adding 'apex/fp16_utils/__init__.py'
    adding 'apex/fp16_utils/fp16_optimizer.py'
    adding 'apex/fp16_utils/fp16util.py'
    adding 'apex/fp16_utils/loss_scaler.py'
    adding 'apex/fused_dense/__init__.py'
    adding 'apex/fused_dense/fused_dense.py'
    adding 'apex/mlp/__init__.py'
    adding 'apex/mlp/mlp.py'
    adding 'apex/multi_tensor_apply/__init__.py'
    adding 'apex/multi_tensor_apply/multi_tensor_apply.py'
    adding 'apex/normalization/__init__.py'
    adding 'apex/normalization/fused_layer_norm.py'
    adding 'apex/optimizers/__init__.py'
    adding 'apex/optimizers/fused_adagrad.py'
    adding 'apex/optimizers/fused_adam.py'
    adding 'apex/optimizers/fused_lamb.py'
    adding 'apex/optimizers/fused_mixed_precision_lamb.py'
    adding 'apex/optimizers/fused_novograd.py'
    adding 'apex/optimizers/fused_sgd.py'

```

```

adding 'apex/parallel/LARC.py'
adding 'apex/parallel/__init__.py'
adding 'apex/parallel/distributed.py'
adding 'apex/parallel/multiproc.py'
adding 'apex/parallel/optimized_sync_batchnorm.py'
adding 'apex/parallel/optimized_sync_batchnorm_kernel.py'
adding 'apex/parallel/sync_batchnorm.py'
adding 'apex/parallel/sync_batchnorm_kernel.py'
adding 'apex/transformer/__init__.py'
adding 'apex/transformer/enums.py'
adding 'apex/transformer/log_util.py'
adding 'apex/transformer/microbatches.py'
adding 'apex/transformer/parallel_state.py'
adding 'apex/transformer/utils.py'
adding 'apex/transformer/_data/__init__.py'
adding 'apex/transformer/_data/_batch_sampler.py'
adding 'apex/transformer/amp/__init__.py'
adding 'apex/transformer/amp/grad_scaler.py'
adding 'apex/transformer/functional/__init__.py'
adding 'apex/transformer/functional/fused_softmax.py'
adding 'apex/transformer/layers/__init__.py'
adding 'apex/transformer/layers/layer_norm.py'
adding 'apex/transformer/pipeline_parallel/__init__.py'
adding 'apex/transformer/pipeline_parallel/_timers.py'
adding 'apex/transformer/pipeline_parallel/p2p_communication.py'
adding 'apex/transformer/pipeline_parallel/utils.py'
adding 'apex/transformer/pipeline_parallel/schedules/__init__.py'
adding 'apex/transformer/pipeline_parallel/schedules/common.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_no_pipelining.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_with_i
nterleaving.py'
adding 'apex/transformer/pipeline_parallel/schedules/fwd_bwd_pipelining_withou
t_interleaving.py'
adding 'apex/transformer/tensor_parallel/__init__.py'
adding 'apex/transformer/tensor_parallel/cross_entropy.py'
adding 'apex/transformer/tensor_parallel/data.py'
adding 'apex/transformer/tensor_parallel/layers.py'
adding 'apex/transformer/tensor_parallel/mappings.py'
adding 'apex/transformer/tensor_parallel/memory.py'
adding 'apex/transformer/tensor_parallel/random.py'
adding 'apex/transformer/tensor_parallel/utils.py'
adding 'apex/transformer/testing/__init__.py'
adding 'apex/transformer/testing/arguments.py'
adding 'apex/transformer/testing/commons.py'
adding 'apex/transformer/testing/distributed_test_base.py'
adding 'apex/transformer/testing/global_vars.py'
adding 'apex/transformer/testing/standalone_bert.py'
adding 'apex/transformer/testing/standalone_gpt.py'

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298244
sha256=bb0104eaef6af8b7216f320f47a46bfcf4491c2e5f4fa157b6a308d52dd66fb3
Stored in directory: /tmp/pip-ephem-wheel-cache-7p22dgpv/wheels/6b/22/a2/16655
26ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

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

```

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 5.1 MB/s
Collecting awscli
  Downloading awscli-1.25.49-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 44.8 MB/s
Collecting botocore==1.27.49
  Downloading botocore-1.27.49-py3-none-any.whl (9.0 MB)
    |                               | 9.0 MB 35.1 MB/s
Collecting s3transfer<0.7.0,>=0.6.0
  Downloading s3transfer-0.6.0-py3-none-any.whl (79 kB)
    |                               | 79 kB 5.7 MB/s
Requirement already satisfied: PyYAML<5.5,>=3.10 in
/usr/local/lib/python3.7/dist-packages (from awscli) (3.13)
Collecting colorama<0.4.5,>=0.2.5
  Downloading colorama-0.4.4-py2.py3-none-any.whl (16 kB)
Collecting rsa<4.8,>=3.1.2
  Downloading rsa-4.7.2-py3-none-any.whl (34 kB)
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 21.9 MB/s
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.49->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==1.27.49->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-1.25.49 botocore-1.27.49 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

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

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting torch==1.7.1
  Downloading torch-1.7.1-cp37-cp37m-manylinux1_x86_64.whl (776.8 MB)
    |                                     | 776.8 MB 16 kB/s
Requirement already satisfied: numpy in /usr/local/lib/python3.7/dist-
packages (from torch==1.7.1) (1.19.2)
Requirement already satisfied: typing-extensions in
/usr/local/lib/python3.7/dist-packages (from torch==1.7.1) (4.1.1)
Installing collected packages: torch
  Attempting uninstall: torch
    Found existing installation: torch 1.4.0

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 1054, done.
remote: Counting objects: 100% (124/124), done.
remote: Compressing objects: 100% (91/91), done.
remote: Total 1054 (delta 51), reused 46 (delta 21), pack-reused 930
Receiving objects: 100% (1054/1054), 257.86 MiB | 30.15 MiB/s, done.
Resolving deltas: 100% (635/635), done.
Checking out files: 100% (1304/1304), done.

```

0.4 remove e move

```
[6]: !ls
```

```

apex          ditto.jpg    LICENSE     requirements.txt  run_all_wdc.py
blocking      ditto_light matcher.py  results_ditto    Snippext_public
configs.json  IC          output      run_all_er_magellan.py train_ditto.py
data         input       README.md   run_all_vary_size.py

```

```

[7]: !rm configs.json
     !mv -n /content/ditto/IC/datasEsErros/FMedium_50_1_2/configs.json /content/
     ↪ditto/

```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

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

Downloading: 100% 433/433 [00:00<00:00, 276kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 354kB/s]

Downloading: 100% 440M/440M [00:06<00:00, 70.2MB/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: 0.5525826811790466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.41379310344827586, 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: 0.48055073618888855
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 2: dev_f1=0.4150943396226415, f1=0.36666666666666664,
best_f1=0.36666666666666664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5682058334350586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.55, f1=0.44897959183673464, best_f1=0.44897959183673464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2739851772785187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.5714285714285714, f1=0.6060606060606061,
best_f1=0.6060606060606061
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2758714258670807
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.5714285714285714,
best_f1=0.5714285714285714
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3109860122203827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8333333333333333, f1=0.7142857142857143,
best_f1=0.7142857142857143
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14555539190769196
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 7: dev_f1=0.8888888888888889, f1=0.6250000000000001,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06927061825990677
/usr/local/lib/python3.7/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=0.88, f1=0.6666666666666667, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009115774184465408
/usr/local/lib/python3.7/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=0.8333333333333333, f1=0.689655172413793,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02912858873605728
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.6666666666666666,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005667021498084068
/usr/local/lib/python3.7/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=0.8, f1=0.689655172413793, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062766363844275475
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.7096774193548386,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005656743887811899
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.7096774193548386,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00964810699224472
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.7096774193548386,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010872578248381615
/usr/local/lib/python3.7/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=0.8148148148148148, f1=0.7096774193548386,
best_f1=0.6250000000000001
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, 140789.99it/s]
/usr/local/lib/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 = 0.8666666666666666
real_f1 = 0.742857142857143
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, 195.73it/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 bert-base-uncased \
--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 bert-base-uncased \
--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: 0.6057059168815613
step: 10, loss: 0.6249390840530396
step: 20, loss: 0.28991952538490295
step: 30, loss: 0.14765149354934692
step: 40, loss: 0.2902633547782898
step: 50, loss: 0.005038924049586058
step: 60, loss: 0.034717198461294174
step: 70, loss: 0.01313649583607912
step: 80, loss: 0.10573048144578934
step: 90, loss: 0.16196073591709137
step: 100, loss: 0.006738121621310711
step: 110, loss: 0.14653350412845612
step: 120, loss: 0.026369348168373108
step: 130, loss: 0.0037982745561748743
step: 140, loss: 0.004080262966454029
step: 150, loss: 0.023118143901228905
step: 160, loss: 0.0037553850561380386
step: 170, loss: 0.08718909323215485
step: 180, loss: 0.005529792048037052
step: 190, loss: 0.003203820902854204
step: 200, loss: 0.004391388967633247
step: 210, loss: 0.005485840141773224
step: 220, loss: 0.010846731252968311
step: 230, loss: 0.04444658011198044
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9681093394077448,
best_f1=0.9681093394077448
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011231360025703907
step: 10, loss: 0.0016194150084629655
step: 20, loss: 0.07608634233474731

```

```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.19865767657756805
step: 40, loss: 0.03128484636545181
step: 50, loss: 0.007728795520961285
step: 60, loss: 0.007805082947015762
step: 70, loss: 0.20388484001159668
step: 80, loss: 0.002426909049972892
step: 90, loss: 0.008195821195840836
step: 100, loss: 0.019979454576969147
step: 110, loss: 0.12144960463047028
step: 120, loss: 0.014371611177921295
step: 130, loss: 0.017746593803167343
step: 140, loss: 0.004192421678453684
step: 150, loss: 0.006578702479600906
step: 160, loss: 0.003945607226341963
step: 170, loss: 0.005896531976759434
step: 180, loss: 0.0016473091673105955
step: 190, loss: 0.0017146453028544784
step: 200, loss: 0.0015600663609802723
step: 210, loss: 0.00038352221599780023
step: 220, loss: 0.09271343797445297
step: 230, loss: 0.023005111142992973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887640449438202, 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: 0.0083121657371521
step: 10, loss: 0.00120362953748554
step: 20, loss: 0.0465824268758297
step: 30, loss: 0.12634100019931793
step: 40, loss: 0.15222521126270294
step: 50, loss: 0.013537850230932236
step: 60, loss: 0.009103160351514816
step: 70, loss: 0.002818470122292638
step: 80, loss: 0.000633150979410857
step: 90, loss: 0.09471793472766876
step: 100, loss: 0.0008514162036590278
step: 110, loss: 0.0029253032989799976
step: 120, loss: 0.01143038272857666
step: 130, loss: 0.010810190811753273

```

```

step: 140, loss: 0.01233943086117506
step: 150, loss: 0.004980152938514948
step: 160, loss: 0.01807701401412487
step: 170, loss: 0.004523350391536951
step: 180, loss: 0.0036915522068738937
step: 190, loss: 0.0691310241818428
step: 200, loss: 0.0032355226576328278
step: 210, loss: 0.002362532541155815
step: 220, loss: 0.0006761363474652171
step: 230, loss: 0.0005713510327041149
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9920903954802259, 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: 0.00154530955478549
step: 10, loss: 0.0006352519849315286
step: 20, loss: 0.002088829642161727
step: 30, loss: 0.0011944733560085297
step: 40, loss: 0.002070953603833914
step: 50, loss: 0.019170274958014488
step: 60, loss: 0.048142191022634506
step: 70, loss: 0.0025307773612439632
step: 80, loss: 0.006638424005359411
step: 90, loss: 0.007923578843474388
step: 100, loss: 0.0010657634120434523
step: 110, loss: 0.0006325837457552552
step: 120, loss: 0.01906394027173519
step: 130, loss: 0.007944339886307716
step: 140, loss: 0.001523657701909542
step: 150, loss: 0.1168130487203598
step: 160, loss: 0.0006912490352988243
step: 170, loss: 0.00301772216334939
step: 180, loss: 0.0003150982956867665
step: 190, loss: 0.00682927155867219
step: 200, loss: 0.00859244167804718
step: 210, loss: 0.0950300320982933
step: 220, loss: 0.03738139942288399
step: 230, loss: 0.0310013797134161
/usr/local/lib/python3.7/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=0.9852440408626559, f1=0.984090909090909,
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: 0.004127363208681345
step: 10, loss: 0.0014917385997250676
step: 20, loss: 0.0023576575331389904
step: 30, loss: 0.0020636345725506544
step: 40, loss: 0.00022460424224846065
step: 50, loss: 0.0008198187570087612
step: 60, loss: 0.0007956565823405981
step: 70, loss: 0.00047991276369430125
step: 80, loss: 0.00036450501647777855
step: 90, loss: 0.00034964020596817136
step: 100, loss: 0.00037072718259878457
step: 110, loss: 0.00012508657528087497
step: 120, loss: 0.00012466726184356958
step: 130, loss: 0.001520377118140459
step: 140, loss: 0.007332806009799242
step: 150, loss: 0.00016348606732208282
step: 160, loss: 0.029484108090400696
step: 170, loss: 0.01515639666467905
step: 180, loss: 0.0011841440573334694
step: 190, loss: 0.00432923948392272
step: 200, loss: 0.11108305305242538
step: 210, loss: 0.0004914669552817941
step: 220, loss: 0.0033489540219306946
step: 230, loss: 0.0002269752585561946
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9751131221719457,
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: 0.0005546605098061264
step: 10, loss: 0.0010469482513144612
step: 20, loss: 0.0014178830897435546
step: 30, loss: 0.0003746331203728914
step: 40, loss: 0.000528948032297194
step: 50, loss: 0.0003120514447800815
step: 60, loss: 0.003598648589104414
step: 70, loss: 0.0006545070209540427
step: 80, loss: 0.0002841294335667044
step: 90, loss: 0.000216399144846946

```

```

step: 100, loss: 0.046288590878248215
step: 110, loss: 0.0022864241618663073
step: 120, loss: 0.0002638286678120494
step: 130, loss: 0.004741283133625984
step: 140, loss: 0.0002561297733336687
step: 150, loss: 0.0006042369641363621
step: 160, loss: 0.0852193683385849
step: 170, loss: 0.0038555297069251537
step: 180, loss: 0.04376240074634552
step: 190, loss: 0.06216825172305107
step: 200, loss: 0.0003576189046725631
step: 210, loss: 0.0005697086453437805
step: 220, loss: 0.00023488979786634445
step: 230, loss: 0.07067608833312988
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.992108229988726, f1=0.9853438556933484,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014946447918191552
step: 10, loss: 0.00012247555423527956
step: 20, loss: 0.0037459498271346092
step: 30, loss: 0.0001590004685567692
step: 40, loss: 8.765744132688269e-05
step: 50, loss: 8.50267824716866e-05
step: 60, loss: 0.006480115465819836
step: 70, loss: 0.0003957234730478376
step: 80, loss: 0.0004370519600342959
step: 90, loss: 3.991100311395712e-05
step: 100, loss: 7.512666343245655e-05
step: 110, loss: 0.0008904412970878184
step: 120, loss: 6.799174298066646e-05
step: 130, loss: 0.0006839009001851082
step: 140, loss: 0.00010598020162433386
step: 150, loss: 0.010932608507573605
step: 160, loss: 0.06640282273292542
step: 170, loss: 0.0003601073694881052
step: 180, loss: 0.002761821262538433
step: 190, loss: 0.00010552666935836896
step: 200, loss: 0.09103026986122131
step: 210, loss: 5.567813423112966e-05

```



```

step: 220, loss: 0.00021817309607286006
step: 230, loss: 8.755864837439731e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9898305084745763, f1=0.9806598407281,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.891528316074982e-05
step: 10, loss: 9.021285950439051e-05
step: 20, loss: 0.00011075612565036863
step: 30, loss: 0.00018946710042655468
step: 40, loss: 0.0006802670541219413
step: 50, loss: 0.0021134186536073685
step: 60, loss: 0.00010406603541923687
step: 70, loss: 0.000690205954015255
step: 80, loss: 0.0016496527241542935
step: 90, loss: 4.2230905819451436e-05
step: 100, loss: 0.00035338691668584943
step: 110, loss: 0.04899315536022186
step: 120, loss: 0.0027607029769569635
step: 130, loss: 0.0017212384846061468
step: 140, loss: 6.996034790063277e-05
step: 150, loss: 0.00038462321390397847
step: 160, loss: 0.0122507493942976
step: 170, loss: 0.000147269107401371
step: 180, loss: 0.00039888842729851604
step: 190, loss: 0.002456110902130604
step: 200, loss: 0.00013678644609171897
step: 210, loss: 0.00020591482461895794
step: 220, loss: 0.0005590745713561773
step: 230, loss: 0.00020331834093667567
/usr/local/lib/python3.7/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=0.9898534385569334, f1=0.9830124575311437,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.786637120763771e-05
step: 10, loss: 0.0012221195502206683
step: 20, loss: 0.002725811442360282
step: 30, loss: 6.084338019718416e-05
step: 40, loss: 0.017018133774399757
step: 50, loss: 4.089051435585134e-05

```

```

step: 60, loss: 6.272690370678902e-05
step: 70, loss: 0.0020300333853811026
step: 80, loss: 0.0003874697722494602
step: 90, loss: 0.0006469295476563275
step: 100, loss: 0.0005936332745477557
step: 110, loss: 3.384244701010175e-05
step: 120, loss: 6.925741763552651e-05
step: 130, loss: 3.042712341994047e-05
step: 140, loss: 0.00014440379163715988
step: 150, loss: 0.00111182383261621
step: 160, loss: 3.2718031434342265e-05
step: 170, loss: 9.163053618976846e-05
step: 180, loss: 0.0048002018593251705
step: 190, loss: 0.00015609081310685724
step: 200, loss: 4.8630405217409134e-05
step: 210, loss: 0.00019308380433358252
step: 220, loss: 3.684161129058339e-05
step: 230, loss: 0.0035415790043771267
/usr/local/lib/python3.7/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=0.9898762654668166, f1=0.9841628959276018,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016358328866772354
step: 10, loss: 3.566339364624582e-05
step: 20, loss: 0.00014491780893877149
step: 30, loss: 0.00016608837177045643
step: 40, loss: 0.00011464438284747303
step: 50, loss: 3.095982174272649e-05
step: 60, loss: 0.005237302277237177
step: 70, loss: 0.004400918260216713
step: 80, loss: 4.6809444029349834e-05
step: 90, loss: 0.0001818091404857114
step: 100, loss: 8.369374700123444e-05
step: 110, loss: 6.977207522140816e-05
step: 120, loss: 0.0001705655304249376
step: 130, loss: 4.132243702770211e-05
step: 140, loss: 0.044467344880104065
step: 150, loss: 0.02464514970779419
step: 160, loss: 5.1789393182843924e-05
step: 170, loss: 0.0002329035778529942
step: 180, loss: 6.553027196787298e-05
step: 190, loss: 0.017616933211684227
step: 200, loss: 4.6820212446618825e-05
step: 210, loss: 0.000251006189500913

```

```

step: 220, loss: 6.502419273601845e-05
step: 230, loss: 0.0003036778944078833
/usr/local/lib/python3.7/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=0.987598647125141, f1=0.975, best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00032149412436410785
step: 10, loss: 5.5901116866152734e-05
step: 20, loss: 0.007959450595080853
step: 30, loss: 0.02049371600151062
step: 40, loss: 0.0007188209565356374
step: 50, loss: 6.258417124627158e-05
step: 60, loss: 0.0004886292736046016
step: 70, loss: 0.0006528326193802059
step: 80, loss: 2.9998607715242542e-05
step: 90, loss: 0.0017944384599104524
step: 100, loss: 5.889625754207373e-05
step: 110, loss: 0.005618568509817123
step: 120, loss: 2.7789817977463827e-05
step: 130, loss: 3.7377325497800484e-05
step: 140, loss: 4.4755637645721436e-05
step: 150, loss: 0.01762053743004799
step: 160, loss: 3.52061542798765e-05
step: 170, loss: 0.025459593161940575
step: 180, loss: 5.503280772245489e-05
step: 190, loss: 3.121280678897165e-05
step: 200, loss: 0.0007312167435884476
step: 210, loss: 2.617705104057677e-05
step: 220, loss: 3.357860987307504e-05
step: 230, loss: 0.0009161244961433113
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9819004524886877,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.372753053554334e-05
step: 10, loss: 0.00014477786317002028
step: 20, loss: 3.8401398342102766e-05
step: 30, loss: 0.0002574701502453536
step: 40, loss: 7.11414177203551e-05
step: 50, loss: 0.0004443468351382762
step: 60, loss: 0.013171347789466381

```

```

step: 70, loss: 8.61724402057007e-05
step: 80, loss: 0.00011765037197619677
step: 90, loss: 3.4043572668451816e-05
step: 100, loss: 3.555212606443092e-05
step: 110, loss: 0.0005288183456286788
step: 120, loss: 0.00011737873137462884
step: 130, loss: 3.101218317169696e-05
step: 140, loss: 0.017637133598327637
step: 150, loss: 4.233918298268691e-05
step: 160, loss: 3.383566581760533e-05
step: 170, loss: 8.337311737705022e-05
step: 180, loss: 0.03416004404425621
step: 190, loss: 0.0004055165045429021
step: 200, loss: 1.8428687326377258e-05
step: 210, loss: 3.19693899655249e-05
step: 220, loss: 6.229530845303088e-05
step: 230, loss: 0.049131207168102264
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9841269841269841,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00015808780153747648
step: 10, loss: 0.008654342964291573
step: 20, loss: 0.00034182227682322264
step: 30, loss: 5.1990635256515816e-05
step: 40, loss: 0.002756014931946993
step: 50, loss: 0.00045580396545119584
step: 60, loss: 0.00018185898079536855
step: 70, loss: 2.2142545276437886e-05
step: 80, loss: 8.097309910226613e-05
step: 90, loss: 0.00010949160059681162
step: 100, loss: 2.761375071713701e-05
step: 110, loss: 0.023051882162690163
step: 120, loss: 0.022922184318304062
step: 130, loss: 3.407763870200142e-05
step: 140, loss: 2.6840038117370568e-05
step: 150, loss: 3.589885091059841e-05
step: 160, loss: 0.0009256793418899179
step: 170, loss: 9.865371976047754e-05
step: 180, loss: 4.8937723477138206e-05
step: 190, loss: 2.0388128177728504e-05
step: 200, loss: 0.00023606230388395488
step: 210, loss: 1.9881208572769538e-05
step: 220, loss: 2.344268432352692e-05

```

```

step: 230, loss: 4.206991434330121e-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=0.9920724801812004, f1=0.9829738933030647,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.3468810468330048e-05
step: 10, loss: 0.011653910391032696
step: 20, loss: 0.002485346281901002
step: 30, loss: 6.027347990311682e-05
step: 40, loss: 0.003054731758311391
step: 50, loss: 3.8196638342924416e-05
step: 60, loss: 3.0240493288147263e-05
step: 70, loss: 4.905212699668482e-05
step: 80, loss: 4.383278064779006e-05
step: 90, loss: 4.4149441237095743e-05
step: 100, loss: 4.526317570707761e-05
step: 110, loss: 0.0004844176000915468
step: 120, loss: 4.6455537813017145e-05
step: 130, loss: 9.166204108623788e-05
step: 140, loss: 8.368404814973474e-05
step: 150, loss: 2.51149704126874e-05
step: 160, loss: 0.014173641800880432
step: 170, loss: 1.5213924598356243e-05
step: 180, loss: 3.394281884538941e-05
step: 190, loss: 0.000622395658865571
step: 200, loss: 2.4217375539592467e-05
step: 210, loss: 4.011925193481147e-05
step: 220, loss: 0.00023634248645976186
step: 230, loss: 0.00010765439947135746
/usr/local/lib/python3.7/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=0.9909706546275394, f1=0.9796839729119639,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.9698776415898465e-05
step: 10, loss: 1.522880484117195e-05
step: 20, loss: 3.686303170979954e-05
step: 30, loss: 2.8128510166425258e-05
step: 40, loss: 6.333569763228297e-05
step: 50, loss: 0.03994724154472351
step: 60, loss: 4.159579475526698e-05

```

```

step: 70, loss: 3.2289273804053664e-05
step: 80, loss: 2.7476553441374563e-05
step: 90, loss: 0.00010964448301820084
step: 100, loss: 2.800216680043377e-05
step: 110, loss: 3.416200343053788e-05
step: 120, loss: 4.098230419913307e-05
step: 130, loss: 0.023003071546554565
step: 140, loss: 2.369965659454465e-05
step: 150, loss: 0.00025210625608451664
step: 160, loss: 2.336433135496918e-05
step: 170, loss: 1.8450991774443537e-05
step: 180, loss: 9.640929783927277e-05
step: 190, loss: 0.0005735497106797993
step: 200, loss: 4.709024142357521e-05
step: 210, loss: 2.6288675144314766e-05
step: 220, loss: 4.059861385030672e-05
step: 230, loss: 2.659029996721074e-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=0.9920724801812004, f1=0.9807037457434733,
best_f1=0.9853438556933484
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
2473it [00:15, 160.74it/s]
load_f1 = 0.9932279909706545
real_f1 = 0.992108229988726
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
4398it [00:24, 182.42it/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 bert-base-uncased \
--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 bert-base-uncased \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

Downloading: 100% 433/433 [00:00<00:00, 490kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 1.23MB/s]
Downloading: 100% 440M/440M [00:13<00:00, 32.9MB/s]
Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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

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

step: 0, loss: 0.6112139225006104
step: 10, loss: 0.5761849880218506
step: 20, loss: 0.4996607005596161
step: 30, loss: 0.09265869855880737
step: 40, loss: 0.13024279475212097
step: 50, loss: 0.13644194602966309
step: 60, loss: 0.05535965785384178
step: 70, loss: 0.12267633527517319
step: 80, loss: 0.04027430713176727
step: 90, loss: 0.3843410015106201


```
step: 100, loss: 0.030194351449608803
step: 110, loss: 0.07422686368227005
step: 120, loss: 0.1277758777141571
step: 130, loss: 0.15399722754955292
step: 140, loss: 0.17723512649536133
step: 150, loss: 0.06345702707767487
step: 160, loss: 0.01942627690732479
step: 170, loss: 0.26898932456970215
step: 180, loss: 0.04384520277380943
step: 190, loss: 0.009539599530398846
step: 200, loss: 0.1833096444606781
step: 210, loss: 0.06889040768146515
step: 220, loss: 0.25833678245544434
step: 230, loss: 0.14003635942935944
step: 240, loss: 0.08440699428319931
step: 250, loss: 0.015164528042078018
step: 260, loss: 0.10453590005636215
step: 270, loss: 0.007850325666368008
step: 280, loss: 0.06237122416496277
step: 290, loss: 0.1456483155488968
step: 300, loss: 0.038811955600976944
step: 310, loss: 0.2561967372894287
step: 320, loss: 0.07986073940992355
step: 330, loss: 0.08630380779504776
step: 340, loss: 0.1208423376083374
step: 350, loss: 0.02691720798611641
step: 360, loss: 0.03054262325167656
step: 370, loss: 0.08073750883340836
step: 380, loss: 0.05517308786511421
step: 390, loss: 0.20947179198265076
step: 400, loss: 0.27701395750045776
step: 410, loss: 0.03809661418199539
step: 420, loss: 0.031067606061697006
step: 430, loss: 0.19620299339294434
step: 440, loss: 0.009951609186828136
step: 450, loss: 0.007634905166924
step: 460, loss: 0.005333814769983292
step: 470, loss: 0.09421862661838531
step: 480, loss: 0.0677407756447792
step: 490, loss: 0.0405094139277935
step: 500, loss: 0.14885996282100677
step: 510, loss: 0.05720456689596176
step: 520, loss: 0.07231215387582779
step: 530, loss: 0.005859813652932644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9310504396112911, f1=0.9259944495837188,

best_f1=0.9259944495837188

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

An input tensor was not cuda.

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

step: 0, loss: 0.141981840133667

step: 10, loss: 0.07232413440942764

step: 20, loss: 0.0224031712859869

step: 30, loss: 0.08984833210706711

step: 40, loss: 0.04487764835357666

step: 50, loss: 0.21352972090244293

step: 60, loss: 0.014138501137495041

step: 70, loss: 0.028903719037771225

step: 80, loss: 0.060970861464738846

step: 90, loss: 0.01994410715997219

step: 100, loss: 0.012419278733432293

step: 110, loss: 0.03905189782381058

step: 120, loss: 0.06537910550832748

step: 130, loss: 0.0625346302986145

step: 140, loss: 0.04938112199306488

step: 150, loss: 0.07006359100341797

step: 160, loss: 0.03457433730363846

step: 170, loss: 0.02773771621286869

step: 180, loss: 0.028659095987677574

step: 190, loss: 0.05763339623808861

step: 200, loss: 0.016431353986263275

step: 210, loss: 0.09919317066669464

step: 220, loss: 0.13107454776763916

step: 230, loss: 0.011858991347253323

step: 240, loss: 0.05689900740981102

step: 250, loss: 0.06110097095370293

step: 260, loss: 0.0016691400669515133

step: 270, loss: 0.27483686804771423

step: 280, loss: 0.02203219011425972

step: 290, loss: 0.0263876561075449

step: 300, loss: 0.17664818465709686

step: 310, loss: 0.007993442006409168

step: 320, loss: 0.12953436374664307

step: 330, loss: 0.028426317498087883

step: 340, loss: 0.011795945465564728

step: 350, loss: 0.003035982372239232

step: 360, loss: 0.034445907920598984

step: 370, loss: 0.1223817765712738

step: 380, loss: 0.02888069674372673

step: 390, loss: 0.06237058341503143

```

step: 400, loss: 0.08173598349094391
step: 410, loss: 0.025683265179395676
step: 420, loss: 0.039788100868463516
step: 430, loss: 0.006068327929824591
step: 440, loss: 0.08100096881389618
step: 450, loss: 0.01855398900806904
step: 460, loss: 0.05996312201023102
step: 470, loss: 0.05318459868431091
step: 480, loss: 0.18412968516349792
step: 490, loss: 0.013300981372594833
step: 500, loss: 0.34288927912712097
step: 510, loss: 0.0510108508169651
step: 520, loss: 0.07898642122745514
step: 530, loss: 0.009891252964735031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9381443298969072, f1=0.9301675977653632,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025555390864610672
step: 10, loss: 0.06895753741264343
step: 20, loss: 0.08603094518184662
step: 30, loss: 0.06780559569597244
step: 40, loss: 0.004980756901204586
step: 50, loss: 0.055251915007829666
step: 60, loss: 0.030890772119164467
step: 70, loss: 0.005589495878666639
step: 80, loss: 0.0011702229967340827
step: 90, loss: 0.009253707714378834
step: 100, loss: 0.13839030265808105
step: 110, loss: 0.009344966150820255
step: 120, loss: 0.03169875219464302
step: 130, loss: 0.009429548867046833
step: 140, loss: 0.006286904215812683
step: 150, loss: 0.017478933557868004
step: 160, loss: 0.020672444254159927
step: 170, loss: 0.04122679680585861
step: 180, loss: 0.014993567019701004
step: 190, loss: 0.002923975232988596
step: 200, loss: 0.022484615445137024
step: 210, loss: 0.0598364882171154

```

```

step: 220, loss: 0.08822594583034515
step: 230, loss: 0.06066465377807617
step: 240, loss: 0.005726207513362169
step: 250, loss: 0.03481512889266014
step: 260, loss: 0.02843300625681877
step: 270, loss: 0.005818720441311598
step: 280, loss: 0.11451701074838638
step: 290, loss: 0.004633158445358276
step: 300, loss: 0.01479597669094801
step: 310, loss: 0.031523630023002625
step: 320, loss: 0.023586519062519073
step: 330, loss: 0.008591785095632076
step: 340, loss: 0.014665552414953709
step: 350, loss: 0.013567444868385792
step: 360, loss: 0.09898122400045395
step: 370, loss: 0.006869898177683353
step: 380, loss: 0.040383633226156235
step: 390, loss: 0.03836481645703316
step: 400, loss: 0.013278207741677761
step: 410, loss: 0.005783507600426674
step: 420, loss: 0.42717841267585754
step: 430, loss: 0.008371870033442974
step: 440, loss: 0.0016620998503640294
step: 450, loss: 0.1072002723813057
step: 460, loss: 0.02065696381032467
step: 470, loss: 0.027829818427562714
step: 480, loss: 0.007155397441238165
step: 490, loss: 0.007289810571819544
step: 500, loss: 0.05713459849357605
step: 510, loss: 0.010871251113712788
step: 520, loss: 0.036254484206438065
step: 530, loss: 0.1272401064634323
/usr/local/lib/python3.7/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=0.9260991580916744, f1=0.9257518796992481,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014480584301054478
step: 10, loss: 0.01570880226790905
step: 20, loss: 0.0008807792910374701
step: 30, loss: 0.01195620372891426
step: 40, loss: 0.003613423788920045
step: 50, loss: 0.004415246658027172
step: 60, loss: 0.009148640558123589
step: 70, loss: 0.0012905426556244493

```

step: 80, loss: 0.029701698571443558
step: 90, loss: 0.12247774749994278
step: 100, loss: 0.006987255532294512
step: 110, loss: 0.015978483483195305
step: 120, loss: 0.00078507763100788
step: 130, loss: 0.0007180646061897278
step: 140, loss: 0.0056296768598258495
step: 150, loss: 0.018458906561136246
step: 160, loss: 0.015689000487327576
step: 170, loss: 0.012396097183227539
step: 180, loss: 0.005440453067421913
step: 190, loss: 0.07545750588178635
step: 200, loss: 0.0035252892412245274
step: 210, loss: 0.02827952615916729
step: 220, loss: 0.0011132658692076802
step: 230, loss: 0.04367717728018761
step: 240, loss: 0.0032131406478583813
step: 250, loss: 0.04976093769073486
step: 260, loss: 0.002312254626303911
step: 270, loss: 0.005895612761378288
step: 280, loss: 0.04715967923402786
step: 290, loss: 0.017230607569217682
step: 300, loss: 0.02445465885102749
step: 310, loss: 0.04071354120969772
step: 320, loss: 0.0940735787153244
step: 330, loss: 0.0838962271809578
step: 340, loss: 0.002465995028614998
step: 350, loss: 0.022304782643914223
step: 360, loss: 0.0714198499917984
step: 370, loss: 0.0038349321112036705
step: 380, loss: 0.0017808957491070032
step: 390, loss: 0.0005586573970504105
step: 400, loss: 0.07651019841432571
step: 410, loss: 0.01435515470802784
step: 420, loss: 0.0074419667944312096
step: 430, loss: 0.17335233092308044
step: 440, loss: 0.02608521282672882
step: 450, loss: 0.0014638088177889585
step: 460, loss: 0.0003831057110801339
step: 470, loss: 0.011126210913062096
step: 480, loss: 0.043039388954639435
step: 490, loss: 0.002927497960627079
step: 500, loss: 0.0587361603975296
step: 510, loss: 0.012861293740570545
step: 520, loss: 0.03427661955356598
step: 530, loss: 0.005666946992278099
/usr/local/lib/python3.7/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=0.9197215777262181, f1=0.9242843951985227,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0309713464230299
step: 10, loss: 0.02404714748263359
step: 20, loss: 0.013725989498198032
step: 30, loss: 0.00954394694417715
step: 40, loss: 0.0006006183102726936
step: 50, loss: 0.00387799134477973
step: 60, loss: 0.1535295695066452
step: 70, loss: 0.004822482820600271
step: 80, loss: 0.0009048462379723787
step: 90, loss: 0.006046958267688751
step: 100, loss: 0.02312321960926056
step: 110, loss: 0.00017570155614521354
step: 120, loss: 0.0006383223808370531
step: 130, loss: 0.0002824361727107316
step: 140, loss: 0.0013256791280582547
step: 150, loss: 0.0007331583765335381
step: 160, loss: 0.0005676455330103636
step: 170, loss: 0.05011812224984169
step: 180, loss: 0.0009732440230436623
step: 190, loss: 0.014293577522039413
step: 200, loss: 0.0011947464663535357
step: 210, loss: 0.0019188327714800835
step: 220, loss: 0.0003954910789616406
step: 230, loss: 0.002116000046953559
step: 240, loss: 0.000468316808110103
step: 250, loss: 0.02699863724410534
step: 260, loss: 0.01286966260522604
step: 270, loss: 0.001986290328204632
step: 280, loss: 0.02963411621749401
step: 290, loss: 0.13662870228290558
step: 300, loss: 0.007060873322188854
step: 310, loss: 0.00795788038522005
step: 320, loss: 0.06898710131645203
step: 330, loss: 0.009382061660289764
step: 340, loss: 0.0027646382804960012
step: 350, loss: 0.0012505385093390942
step: 360, loss: 0.009228934533894062
step: 370, loss: 0.03927990049123764
step: 380, loss: 0.0025073334109038115
step: 390, loss: 0.00022539231576956809
step: 400, loss: 0.01904618926346302
step: 410, loss: 0.003931358456611633
```

```

step: 420, loss: 0.014294499531388283
step: 430, loss: 0.0009824020089581609
step: 440, loss: 0.10423880070447922
step: 450, loss: 0.001957148080691695
step: 460, loss: 0.0008735321462154388
step: 470, loss: 0.017923589795827866
step: 480, loss: 0.017003824934363365
step: 490, loss: 0.0006280196830630302
step: 500, loss: 0.0017913796473294497
step: 510, loss: 0.08773525804281235
step: 520, loss: 0.01060349028557539
step: 530, loss: 0.00016940332716330886
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.9227941176470589, f1=0.9227230910763569,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.055101603269577026
step: 10, loss: 0.0009454225655645132
step: 20, loss: 0.13402430713176727
step: 30, loss: 0.0024280562065541744
step: 40, loss: 0.011862417683005333
step: 50, loss: 0.047061316668987274
step: 60, loss: 0.008964655920863152
step: 70, loss: 0.0030879685655236244
step: 80, loss: 0.005744440481066704
step: 90, loss: 0.0028458177112042904
step: 100, loss: 0.00030726855038665235
step: 110, loss: 0.0009588910033926368
step: 120, loss: 0.00242018373683095
step: 130, loss: 0.0015287597198039293
step: 140, loss: 0.000966014398727566
step: 150, loss: 0.0004241330025251955
step: 160, loss: 0.0002569913922343403
step: 170, loss: 0.005830035079270601
step: 180, loss: 0.00039082023431546986
step: 190, loss: 0.0679408386349678
step: 200, loss: 0.00024391130136791617
step: 210, loss: 0.0005913501954637468
step: 220, loss: 0.000399152806494385
step: 230, loss: 0.0005877743824385107
step: 240, loss: 0.0002456612419337034
step: 250, loss: 0.05057136341929436
step: 260, loss: 9.742511610966176e-05

```

```

step: 270, loss: 0.0009171712445095181
step: 280, loss: 0.0004912580479867756
step: 290, loss: 0.0001287950435653329
step: 300, loss: 0.001994804944843054
step: 310, loss: 0.010777272284030914
step: 320, loss: 0.000281800574157387
step: 330, loss: 0.0010980108054354787
step: 340, loss: 0.08408397436141968
step: 350, loss: 0.0006561963236890733
step: 360, loss: 0.01449219137430191
step: 370, loss: 0.17815320193767548
step: 380, loss: 0.00039543944876641035
step: 390, loss: 0.002338081831112504
step: 400, loss: 0.00664733350276947
step: 410, loss: 0.006909026298671961
step: 420, loss: 0.0029447844717651606
step: 430, loss: 0.0008453644113615155
step: 440, loss: 0.0001506960397819057
step: 450, loss: 0.0002687560918275267
step: 460, loss: 0.0005530666094273329
step: 470, loss: 0.0021866776514798403
step: 480, loss: 0.00981609895825386
step: 490, loss: 0.008113800548017025
step: 500, loss: 0.0016185024287551641
step: 510, loss: 0.011861443519592285
step: 520, loss: 0.001848089974373579
step: 530, loss: 0.023339437320828438
/usr/local/lib/python3.7/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=0.9206349206349207, f1=0.9171736078614879,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004695162642747164
step: 10, loss: 0.009927350096404552
step: 20, loss: 0.012538718059659004
step: 30, loss: 0.0022897240705788136
step: 40, loss: 0.001601825701072812
step: 50, loss: 0.0007968173595145345
step: 60, loss: 0.004262837581336498
step: 70, loss: 0.026606515049934387
step: 80, loss: 0.0033444419968873262
step: 90, loss: 0.00044489200809039176
step: 100, loss: 7.559629011666402e-05
step: 110, loss: 0.001394249964505434
step: 120, loss: 0.009921983815729618

```



```

step: 130, loss: 0.004778003785759211
step: 140, loss: 0.0001034564629662782
step: 150, loss: 0.00048337061889469624
step: 160, loss: 0.00010973892494803295
step: 170, loss: 0.00012068434443790466
step: 180, loss: 0.001365081174299121
step: 190, loss: 0.004476046189665794
step: 200, loss: 0.00014024759002495557
step: 210, loss: 0.0004119114892091602
step: 220, loss: 0.00020024435070808977
step: 230, loss: 0.10160300880670547
step: 240, loss: 0.0046157557517290115
step: 250, loss: 0.0007817660225555301
step: 260, loss: 0.00016086321556940675
step: 270, loss: 0.0011260728351771832
step: 280, loss: 0.0012583109783008695
step: 290, loss: 9.229940042132512e-05
step: 300, loss: 0.0035436474718153477
step: 310, loss: 0.001122034853324294
step: 320, loss: 0.0012480239383876324
step: 330, loss: 0.0009961823234334588
step: 340, loss: 0.00794603955000639
step: 350, loss: 0.0006650958093814552
step: 360, loss: 0.0022141719236969948
step: 370, loss: 0.00031907297670841217
step: 380, loss: 0.00017419569485355169
step: 390, loss: 0.000146760226925835
step: 400, loss: 0.008396129123866558
step: 410, loss: 0.002393783302977681
step: 420, loss: 0.010930193588137627
step: 430, loss: 0.00020432836026884615
step: 440, loss: 0.0019443127093836665
step: 450, loss: 0.007447691168636084
step: 460, loss: 0.0015222755027934909
step: 470, loss: 0.0014924912247806787
step: 480, loss: 0.11942704021930695
step: 490, loss: 0.0012120319297537208
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 500, loss: 0.002719215117394924
step: 510, loss: 0.0037601543590426445
step: 520, loss: 0.018899012356996536
step: 530, loss: 0.036000411957502365
/usr/local/lib/python3.7/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=0.9271523178807948, f1=0.9201700519603213,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003187562106177211
step: 10, loss: 0.000205141375772655
step: 20, loss: 9.316526120528579e-05
step: 30, loss: 0.0004025712551083416
step: 40, loss: 0.004275472369045019
step: 50, loss: 0.04144607484340668
step: 60, loss: 0.004131444729864597
step: 70, loss: 0.00011599927529459819
step: 80, loss: 0.0010925009846687317
step: 90, loss: 0.00018968664517160505
step: 100, loss: 8.836847700877115e-05
step: 110, loss: 5.848565342603251e-05
step: 120, loss: 0.003974487539380789
step: 130, loss: 0.004000482149422169
step: 140, loss: 0.0038218870759010315
step: 150, loss: 0.003543214173987508
step: 160, loss: 0.009168018586933613
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.00138648203574121
step: 180, loss: 6.692620081594214e-05
step: 190, loss: 0.00268665817566216
step: 200, loss: 0.0007229113834910095
step: 210, loss: 0.0007010906119830906
step: 220, loss: 0.0029607012402266264
step: 230, loss: 0.00015119636373128742
step: 240, loss: 0.000309529947116971
step: 250, loss: 6.677309283986688e-05
step: 260, loss: 0.00010210338223259896
step: 270, loss: 0.08152545243501663
step: 280, loss: 0.04011954367160797
step: 290, loss: 0.0008828580612316728
step: 300, loss: 0.0017619171412661672
step: 310, loss: 0.00044675092794932425
step: 320, loss: 0.002345845801755786
step: 330, loss: 0.007890750654041767
step: 340, loss: 0.000564802554436028
step: 350, loss: 0.004381836857646704
step: 360, loss: 0.0002429976884741336
step: 370, loss: 4.035781603306532e-05
step: 380, loss: 0.009253866970539093
step: 390, loss: 0.16533003747463226
step: 400, loss: 0.008340463042259216
step: 410, loss: 7.015430310275406e-05
step: 420, loss: 0.015119336545467377
step: 430, loss: 0.013087012805044651
step: 440, loss: 0.0029018244240432978
```

```

step: 450, loss: 0.00013819111336488277
step: 460, loss: 0.00035257430863566697
step: 470, loss: 0.02077815867960453
step: 480, loss: 0.00012171010894235224
step: 490, loss: 0.004175127483904362
step: 500, loss: 0.00019030302064493299
step: 510, loss: 0.00023389095440506935
step: 520, loss: 0.00328422081656754
step: 530, loss: 7.562293467344716e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9304964539007093, f1=0.9234382339126351,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018530647503212094
step: 10, loss: 9.378222603118047e-05
step: 20, loss: 0.0006397615070454776
step: 30, loss: 0.000168783706612885
step: 40, loss: 0.00027663842774927616
step: 50, loss: 8.079451799858361e-05
step: 60, loss: 0.00010421354090794921
step: 70, loss: 0.0026718818116933107
step: 80, loss: 0.00039502454455941916
step: 90, loss: 0.0008229714003391564
step: 100, loss: 0.00016831573157105595
step: 110, loss: 0.00011153090599691495
step: 120, loss: 0.0002445206919219345
step: 130, loss: 0.03248611465096474
step: 140, loss: 0.00010292528168065473
step: 150, loss: 5.0329239456914365e-05
step: 160, loss: 0.0011988675687462091
step: 170, loss: 0.00679492624476552
step: 180, loss: 9.072176908375695e-05
step: 190, loss: 8.56980259413831e-05
step: 200, loss: 0.00011449181329226121
step: 210, loss: 0.00018526165513321757
step: 220, loss: 0.0028199797961860895
step: 230, loss: 4.249593257554807e-05
step: 240, loss: 0.04522290453314781
step: 250, loss: 0.007519814185798168
step: 260, loss: 0.010651445016264915
step: 270, loss: 5.440631139208563e-05
step: 280, loss: 0.0866638720035553
step: 290, loss: 0.0062584057450294495
step: 300, loss: 6.24194581178017e-05

```

```

step: 310, loss: 0.009206963703036308
step: 320, loss: 0.0005285089137032628
step: 330, loss: 0.014420008286833763
step: 340, loss: 0.0010769875952973962
step: 350, loss: 0.0010720244608819485
step: 360, loss: 5.935695662628859e-05
step: 370, loss: 0.004094521515071392
step: 380, loss: 0.00010272636427544057
step: 390, loss: 6.621023931074888e-05
step: 400, loss: 0.009515371173620224
step: 410, loss: 0.00014990394993219525
step: 420, loss: 4.746398553834297e-05
step: 430, loss: 0.0006556441658176482
step: 440, loss: 0.0030421356204897165
step: 450, loss: 8.290470577776432e-05
step: 460, loss: 0.0013807103969156742
step: 470, loss: 2.9533362976508215e-05
step: 480, loss: 0.00010115477198269218
step: 490, loss: 0.001002933015115559
step: 500, loss: 0.00042985714389942586
step: 510, loss: 0.005186126101762056
step: 520, loss: 0.0002182797179557383
step: 530, loss: 7.715677202213556e-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=0.927348449791763, f1=0.9271889400921659,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016552578017581254
step: 10, loss: 3.618170376284979e-05
step: 20, loss: 5.45143338968046e-05
step: 30, loss: 8.025570423342288e-05
step: 40, loss: 0.0015513449907302856
step: 50, loss: 0.0002445452264510095
step: 60, loss: 6.410440983017907e-05
step: 70, loss: 7.009419641690329e-05
step: 80, loss: 3.289704545750283e-05
step: 90, loss: 7.993663166416809e-05
step: 100, loss: 7.07968429196626e-05
step: 110, loss: 8.732483547646552e-05
step: 120, loss: 8.596004045102745e-05
step: 130, loss: 0.00018424341396894306
step: 140, loss: 0.00016023797797970474
step: 150, loss: 5.479163155541755e-05
step: 160, loss: 4.714962051366456e-05

```

```

step: 170, loss: 5.3077492339070886e-05
step: 180, loss: 0.0003256300697103143
step: 190, loss: 0.00023009728465694934
step: 200, loss: 0.00012197933392599225
step: 210, loss: 0.0017719839233905077
step: 220, loss: 0.005943921394646168
step: 230, loss: 0.0003646414843387902
step: 240, loss: 0.00010495405149413273
step: 250, loss: 0.01319038961082697
step: 260, loss: 0.0025964598171412945
step: 270, loss: 0.0007657698588445783
step: 280, loss: 6.903061148477718e-05
step: 290, loss: 8.918100502341986e-05
step: 300, loss: 2.994639908138197e-05
step: 310, loss: 0.00013613175542559475
step: 320, loss: 0.001313585788011551
step: 330, loss: 3.666717020678334e-05
step: 340, loss: 0.0021475041285157204
step: 350, loss: 6.923313048901036e-05
step: 360, loss: 0.00022033015557099134
step: 370, loss: 0.000682466255966574
step: 380, loss: 0.0006798045942559838
step: 390, loss: 0.0007704552263021469
step: 400, loss: 0.00019797119603026658
step: 410, loss: 0.001344421529211104
step: 420, loss: 3.590175765566528e-05
step: 430, loss: 2.776286783046089e-05
step: 440, loss: 0.00029067654395475984
step: 450, loss: 4.597817314788699e-05
step: 460, loss: 4.1038412746274844e-05
step: 470, loss: 0.00133479421492666
step: 480, loss: 3.269232911407016e-05
step: 490, loss: 6.364154251059517e-05
step: 500, loss: 0.004135034047067165
step: 510, loss: 8.991961658466607e-05
step: 520, loss: 0.0002095382078550756
step: 530, loss: 0.0007657863316126168
/usr/local/lib/python3.7/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=0.9280677009873061, f1=0.924092409240924,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004474273882806301
step: 10, loss: 0.00017074205970857292
step: 20, loss: 0.00010291920625604689

```

step: 30, loss: 0.0008478441741317511
step: 40, loss: 0.00036078336415812373
step: 50, loss: 0.000507600954733789
step: 60, loss: 0.004728947300463915
step: 70, loss: 7.565146370325238e-05
step: 80, loss: 0.00028225427377037704
step: 90, loss: 6.374937947839499e-05
step: 100, loss: 0.00026159427943639457
step: 110, loss: 0.00016110032447613776
step: 120, loss: 4.6159686462488025e-05
step: 130, loss: 4.396999065647833e-05
step: 140, loss: 9.766151197254658e-05
step: 150, loss: 3.933841435355134e-05
step: 160, loss: 4.39617142546922e-05
step: 170, loss: 0.00011027992877643555
step: 180, loss: 0.0014293218264356256
step: 190, loss: 0.003264656523242593
step: 200, loss: 0.005125797353684902
step: 210, loss: 0.00044947597780264914
step: 220, loss: 0.00019214858184568584
step: 230, loss: 0.0020950704347342253
step: 240, loss: 0.00023266905918717384
step: 250, loss: 2.0503368432400748e-05
step: 260, loss: 0.0021576951257884502
step: 270, loss: 5.4526452004211023e-05
step: 280, loss: 4.114330295124091e-05
step: 290, loss: 0.00017048648442141712
step: 300, loss: 3.6687841202365234e-05
step: 310, loss: 0.00010200667747994885
step: 320, loss: 0.00019051373237743974
step: 330, loss: 0.0006264232797548175
step: 340, loss: 0.0005371414008550346
step: 350, loss: 4.280672874301672e-05
step: 360, loss: 3.8507165299961343e-05
step: 370, loss: 0.0012326156720519066
step: 380, loss: 3.912351530743763e-05
step: 390, loss: 0.00010664344154065475
step: 400, loss: 2.4131701138685457e-05
step: 410, loss: 2.1952711904305033e-05
step: 420, loss: 7.203753921203315e-05
step: 430, loss: 2.3077669538906775e-05
step: 440, loss: 0.0006725882994942367
step: 450, loss: 3.556214505806565e-05
step: 460, loss: 0.0020412991289049387
step: 470, loss: 2.949561530840583e-05
step: 480, loss: 7.251830538734794e-05
step: 490, loss: 0.0030329707078635693
step: 500, loss: 0.0011423625983297825

```

step: 510, loss: 0.0002847050200216472
step: 520, loss: 5.1397997594904155e-05
step: 530, loss: 0.0018529309891164303
/usr/local/lib/python3.7/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=0.9263256687001408, f1=0.9258215962441314,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001378189626848325
step: 10, loss: 4.6443954488495365e-05
step: 20, loss: 0.006621797569096088
step: 30, loss: 7.466043462045491e-05
step: 40, loss: 0.0009522641776129603
step: 50, loss: 0.004660172387957573
step: 60, loss: 0.001177601283416152
step: 70, loss: 8.177241397788748e-05
step: 80, loss: 0.008519290946424007
step: 90, loss: 0.0001515152835054323
step: 100, loss: 0.00033832225017249584
step: 110, loss: 6.465181650128216e-05
step: 120, loss: 3.264708357164636e-05
step: 130, loss: 0.0011853532632812858
step: 140, loss: 9.842684812610969e-05
step: 150, loss: 2.119281271006912e-05
step: 160, loss: 3.1804673199076205e-05
step: 170, loss: 2.978267366415821e-05
step: 180, loss: 4.3723921407945454e-05
step: 190, loss: 3.732811092049815e-05
step: 200, loss: 0.00014235172420740128
step: 210, loss: 0.018962424248456955
step: 220, loss: 4.188150342088193e-05
step: 230, loss: 2.130814755219035e-05
step: 240, loss: 0.0002943847212009132
step: 250, loss: 2.6906895072897896e-05
step: 260, loss: 0.0016019648173823953
step: 270, loss: 5.462495028041303e-05
step: 280, loss: 0.0007261126884259284
step: 290, loss: 0.0001615135552128777
step: 300, loss: 0.00010822966578416526
step: 310, loss: 0.00020385670359246433
step: 320, loss: 2.2589138097828254e-05
step: 330, loss: 0.0008751269779168069
step: 340, loss: 3.0378401788766496e-05
step: 350, loss: 5.815256736241281e-05
step: 360, loss: 0.012267555110156536

```

```

step: 370, loss: 3.745492722373456e-05
step: 380, loss: 9.121902985498309e-05
step: 390, loss: 0.0009852616349235177
step: 400, loss: 0.004419515375047922
step: 410, loss: 0.00010554204345680773
step: 420, loss: 0.002274012193083763
step: 430, loss: 3.911148814950138e-05
step: 440, loss: 0.0001002444259938784
step: 450, loss: 5.17457774549257e-05
step: 460, loss: 0.0004456797323655337
step: 470, loss: 6.288695294642821e-05
step: 480, loss: 0.00012362045526970178
step: 490, loss: 0.0010467636166140437
step: 500, loss: 3.0072797017055564e-05
step: 510, loss: 0.002017635153606534
step: 520, loss: 0.000219856490730308
step: 530, loss: 0.0036740568466484547
/usr/local/lib/python3.7/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=0.923728813559322, f1=0.9177693761814745,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006747301667928696
step: 10, loss: 7.01288809068501e-05
step: 20, loss: 5.972134385956451e-05
step: 30, loss: 3.495650889817625e-05
step: 40, loss: 3.8159701944096014e-05
step: 50, loss: 4.7122641262831166e-05
step: 60, loss: 8.649878873256966e-05
step: 70, loss: 0.00020153749210294336
step: 80, loss: 3.603811637731269e-05
step: 90, loss: 0.0013704978628084064
step: 100, loss: 3.146562448819168e-05
step: 110, loss: 2.7912423320231028e-05
step: 120, loss: 0.00011165066825924441
step: 130, loss: 2.2619406081503257e-05
step: 140, loss: 0.00022991005971562117
step: 150, loss: 0.00013647427840624005
step: 160, loss: 2.2146034098113887e-05
step: 170, loss: 4.880688356934115e-05
step: 180, loss: 0.0002629792725201696
step: 190, loss: 0.0003456133999861777
step: 200, loss: 2.6385348974145018e-05
step: 210, loss: 4.47713682660833e-05
step: 220, loss: 0.0004976309137418866

```



```

step: 230, loss: 2.8312750146142207e-05
step: 240, loss: 0.006287263706326485
step: 250, loss: 2.199730442953296e-05
step: 260, loss: 2.194490298279561e-05
step: 270, loss: 2.1560786990448833e-05
step: 280, loss: 2.0559507902362384e-05
step: 290, loss: 2.37437998293899e-05
step: 300, loss: 0.00013353521353565156
step: 310, loss: 0.005166496615856886
step: 320, loss: 0.0022634451743215322
step: 330, loss: 0.00120248319581151
step: 340, loss: 3.072402978432365e-05
step: 350, loss: 3.25188048009295e-05
step: 360, loss: 2.385589868936222e-05
step: 370, loss: 6.895154365338385e-05
step: 380, loss: 5.355978282750584e-05
step: 390, loss: 0.013077208772301674
step: 400, loss: 0.0036656649317592382
step: 410, loss: 0.00032640955760143697
step: 420, loss: 1.800407153496053e-05
step: 430, loss: 0.014620697125792503
step: 440, loss: 0.00017805893730837852
step: 450, loss: 9.40446843742393e-05
step: 460, loss: 0.0005873360205441713
step: 470, loss: 7.810159877408296e-05
step: 480, loss: 0.001256524003110826
step: 490, loss: 0.000527491036336869
step: 500, loss: 2.9155949960113503e-05
step: 510, loss: 0.00022757890110369772
step: 520, loss: 0.0042352089658379555
step: 530, loss: 1.8164069842896424e-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=0.9254571026722925, f1=0.9255419415645617,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.246965494938195e-05
step: 10, loss: 1.1671240827126894e-05
step: 20, loss: 2.496096567483619e-05
step: 30, loss: 4.8196041461778805e-05
step: 40, loss: 3.375884625711478e-05
step: 50, loss: 0.00012317353684920818
step: 60, loss: 1.1593017006816808e-05
step: 70, loss: 3.514117997838184e-05
step: 80, loss: 1.574660745973233e-05

```

```
step: 90, loss: 3.819811172434129e-05
step: 100, loss: 0.0004016718012280762
step: 110, loss: 2.714803304115776e-05
step: 120, loss: 4.0968225221149623e-05
step: 130, loss: 0.0003177060862071812
step: 140, loss: 3.013790410477668e-05
step: 150, loss: 3.348951577208936e-05
step: 160, loss: 0.00030003051506355405
step: 170, loss: 2.3099615646060556e-05
step: 180, loss: 1.50052692333702e-05
step: 190, loss: 0.00025733900838531554
step: 200, loss: 2.7375761419534683e-05
step: 210, loss: 4.68579055450391e-05
step: 220, loss: 1.5146812074817717e-05
step: 230, loss: 2.382268212386407e-05
step: 240, loss: 6.968898378545418e-05
step: 250, loss: 3.254602779634297e-05
step: 260, loss: 2.222808325313963e-05
step: 270, loss: 3.84837003366556e-05
step: 280, loss: 1.4889814337948337e-05
step: 290, loss: 0.008959325961768627
step: 300, loss: 4.7663656005170196e-05
step: 310, loss: 0.006964335683733225
step: 320, loss: 0.00046088328235782683
step: 330, loss: 5.382368544815108e-05
step: 340, loss: 7.563051622128114e-05
step: 350, loss: 0.0001143874615081586
step: 360, loss: 0.0366409495472908
step: 370, loss: 3.0279197744675912e-05
step: 380, loss: 5.1900860853493214e-05
step: 390, loss: 0.006678550038486719
step: 400, loss: 0.0030640612822026014
step: 410, loss: 2.6876259653363377e-05
step: 420, loss: 0.00925316009670496
step: 430, loss: 0.00015686865663155913
step: 440, loss: 6.515980930998921e-05
step: 450, loss: 7.254620140884072e-05
step: 460, loss: 4.306982009438798e-05
step: 470, loss: 0.00020247881184332073
step: 480, loss: 1.817515294533223e-05
step: 490, loss: 1.2393935321597382e-05
step: 500, loss: 1.6104178939713165e-05
step: 510, loss: 4.070586510351859e-05
step: 520, loss: 1.592160333530046e-05
step: 530, loss: 7.49628379708156e-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=0.9240093240093239, f1=0.9256661991584852,
best_f1=0.9301675977653632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
 warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021986755018588156
step: 10, loss: 1.728108327370137e-05
step: 20, loss: 1.9903836800949648e-05
step: 30, loss: 0.0003042147436644882
step: 40, loss: 0.0387883298099041
step: 50, loss: 4.194149005343206e-05
step: 60, loss: 2.2361076844390482e-05
step: 70, loss: 0.0035732053220272064
step: 80, loss: 2.8835651392000727e-05
step: 90, loss: 1.5500758308917284e-05
step: 100, loss: 9.497452265350148e-05
step: 110, loss: 4.304706089897081e-05
step: 120, loss: 4.5328251871978864e-05
step: 130, loss: 0.023808741942048073
step: 140, loss: 1.5221266949083656e-05
step: 150, loss: 2.691666486498434e-05
step: 160, loss: 5.575110481004231e-05
step: 170, loss: 5.871309622307308e-05
step: 180, loss: 1.3001111256016884e-05
step: 190, loss: 3.092388215009123e-05
step: 200, loss: 2.7014413717552088e-05
step: 210, loss: 0.0002583371242508292
step: 220, loss: 2.705748556763865e-05
step: 230, loss: 0.0016530497232452035
step: 240, loss: 2.87727452814579e-05
step: 250, loss: 1.614516440895386e-05
step: 260, loss: 4.093044844921678e-05
step: 270, loss: 1.4260249372455291e-05
step: 280, loss: 1.6543779565836303e-05
step: 290, loss: 1.100070949178189e-05
step: 300, loss: 1.247959335159976e-05
step: 310, loss: 2.6715410058386624e-05
step: 320, loss: 1.235294621437788e-05
step: 330, loss: 2.1345151253626682e-05
step: 340, loss: 1.7232843674719334e-05
step: 350, loss: 1.8688608179218136e-05
step: 360, loss: 0.00021546060452237725
step: 370, loss: 1.1958080904150847e-05
step: 380, loss: 1.670770870987326e-05
step: 390, loss: 2.620868690428324e-05
step: 400, loss: 6.020576984155923e-05
step: 410, loss: 3.328904858790338e-05
step: 420, loss: 2.5215103960363194e-05

```

step: 430, loss: 7.446280505973846e-05
step: 440, loss: 2.060402766801417e-05
step: 450, loss: 1.2449797395674977e-05
step: 460, loss: 1.744522523949854e-05
step: 470, loss: 0.009204726666212082
step: 480, loss: 1.2211364264658187e-05
step: 490, loss: 1.1037958756787702e-05
step: 500, loss: 0.00010896598541876301
step: 510, loss: 1.535172668809537e-05
step: 520, loss: 1.591052750882227e-05
step: 530, loss: 4.3820979044539854e-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=0.9250116441546343, f1=0.9245107176141659,
best_f1=0.9301675977653632
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:23, 247.79it/s]
load_f1 = 0.933826931975937

```

```

real_f1 = 0.933579335793358
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 [00:18, 242.56it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5598036050796509
step: 10, loss: 0.3724798858165741
step: 20, loss: 0.3711826205253601
step: 30, loss: 0.32394054532051086
step: 40, loss: 0.15782466530799866
step: 50, loss: 0.4257924556732178
step: 60, loss: 0.2819114923477173
step: 70, loss: 0.1687166690826416
step: 80, loss: 0.19220884144306183
step: 90, loss: 0.32432931661605835
step: 100, loss: 0.29886823892593384
step: 110, loss: 0.2283511608839035
step: 120, loss: 0.20441561937332153
step: 130, loss: 0.1765952706336975
step: 140, loss: 0.2474832385778427
step: 150, loss: 0.24614039063453674
step: 160, loss: 0.23412856459617615
step: 170, loss: 0.2832360863685608
step: 180, loss: 0.10937520861625671

```

```

step: 190, loss: 0.23019027709960938
step: 200, loss: 0.2768842875957489
step: 210, loss: 0.1954365372657776
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5577689243027888, f1=0.6029106029106029,
best_f1=0.6029106029106029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1887277513742447
step: 10, loss: 0.2917536199092865
step: 20, loss: 0.1877974420785904
step: 30, loss: 0.2271943837404251
step: 40, loss: 0.3054088354110718
step: 50, loss: 0.13405056297779083
step: 60, loss: 0.312446653842926
step: 70, loss: 0.15952517092227936
step: 80, loss: 0.15314051508903503
step: 90, loss: 0.11414064466953278
step: 100, loss: 0.016651280224323273
step: 110, loss: 0.1845061033964157
step: 120, loss: 0.14541210234165192
step: 130, loss: 0.03561607748270035
step: 140, loss: 0.14115968346595764
step: 150, loss: 0.2504834532737732
step: 160, loss: 0.18280285596847534
step: 170, loss: 0.10971717536449432
step: 180, loss: 0.18671655654907227
step: 190, loss: 0.27811598777770996
step: 200, loss: 0.07039676606655121
step: 210, loss: 0.14356347918510437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5807692307692308, f1=0.5964912280701755,
best_f1=0.5964912280701755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05592883750796318
step: 10, loss: 0.1395963430404663
step: 20, loss: 0.1440947949886322
step: 30, loss: 0.06494157016277313
step: 40, loss: 0.1443946808576584
step: 50, loss: 0.10042446106672287
step: 60, loss: 0.15180723369121552
step: 70, loss: 0.07649175077676773
step: 80, loss: 0.19419313967227936
step: 90, loss: 0.05536219850182533
step: 100, loss: 0.1487531065940857
step: 110, loss: 0.26447027921676636
step: 120, loss: 0.14968241751194
step: 130, loss: 0.2001883089542389
step: 140, loss: 0.1021946594119072
step: 150, loss: 0.1743331253528595
step: 160, loss: 0.03898724168539047
step: 170, loss: 0.2157101333141327
step: 180, loss: 0.09947267919778824
step: 190, loss: 0.188408762216568
step: 200, loss: 0.0311172716319561
step: 210, loss: 0.13408252596855164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5841392649903289, f1=0.6330097087378641,
best_f1=0.6330097087378641
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1761915683746338
step: 10, loss: 0.07003211975097656
step: 20, loss: 0.041822370141744614
step: 30, loss: 0.12127557396888733
step: 40, loss: 0.018630629405379295
step: 50, loss: 0.3050701320171356
step: 60, loss: 0.19458192586898804
step: 70, loss: 0.22916020452976227
step: 80, loss: 0.08853716403245926
step: 90, loss: 0.05604640394449234
step: 100, loss: 0.23611435294151306
step: 110, loss: 0.09689179807901382
step: 120, loss: 0.14567790925502777

```



```

step: 130, loss: 0.21805283427238464
step: 140, loss: 0.19200602173805237
step: 150, loss: 0.03883733972907066
step: 160, loss: 0.05303743854165077
step: 170, loss: 0.15159285068511963
step: 180, loss: 0.3295295238494873
step: 190, loss: 0.06136636435985565
step: 200, loss: 0.15663041174411774
step: 210, loss: 0.19279645383358002
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6131386861313869, f1=0.6233766233766235,
best_f1=0.6233766233766235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1358800083398819
step: 10, loss: 0.3075452148914337
step: 20, loss: 0.37591758370399475
step: 30, loss: 0.08910229057073593
step: 40, loss: 0.0341729111969471
step: 50, loss: 0.06093299388885498
step: 60, loss: 0.09123119711875916
step: 70, loss: 0.10841576755046844
step: 80, loss: 0.08981432020664215
step: 90, loss: 0.1467059850692749
step: 100, loss: 0.0020013817120343447
step: 110, loss: 0.1012880727648735
step: 120, loss: 0.22137020528316498
step: 130, loss: 0.08975016325712204
step: 140, loss: 0.10814248770475388
step: 150, loss: 0.06598053127527237
step: 160, loss: 0.2073296308517456
step: 170, loss: 0.06348360329866409
step: 180, loss: 0.03568857163190842
step: 190, loss: 0.050461504608392715
step: 200, loss: 0.08577824383974075
step: 210, loss: 0.05240341275930405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6134453781512605, f1=0.6150627615062761,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06791884452104568
step: 10, loss: 0.056246429681777954
step: 20, loss: 0.013278260827064514
step: 30, loss: 0.005311715882271528
step: 40, loss: 0.03950103372335434
step: 50, loss: 0.035311561077833176
step: 60, loss: 0.12717430293560028
step: 70, loss: 0.03900650888681412
step: 80, loss: 0.2401895523071289
step: 90, loss: 0.1556638926267624
step: 100, loss: 0.0034073875285685062
step: 110, loss: 0.05151527374982834
step: 120, loss: 0.12677089869976044
step: 130, loss: 0.35700175166130066
step: 140, loss: 0.09051655232906342
step: 150, loss: 0.044989585876464844
step: 160, loss: 0.056572359055280685
step: 170, loss: 0.17216870188713074
step: 180, loss: 0.04845832288265228
step: 190, loss: 0.06726638972759247
step: 200, loss: 0.01941925659775734
step: 210, loss: 0.03244544565677643
/usr/local/lib/python3.7/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=0.5783132530120482, f1=0.6223091976516635,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04500409960746765
step: 10, loss: 0.18394100666046143
step: 20, loss: 0.012024104595184326
step: 30, loss: 0.044502127915620804
step: 40, loss: 0.021453870460391045
step: 50, loss: 0.1144186332821846
step: 60, loss: 0.04227632284164429
step: 70, loss: 0.01750374399125576
step: 80, loss: 0.10415192693471909
step: 90, loss: 0.03452767804265022
step: 100, loss: 0.051370769739151

```

```

step: 110, loss: 0.11434818804264069
step: 120, loss: 0.04931601881980896
step: 130, loss: 0.1058887243270874
step: 140, loss: 0.021735114976763725
step: 150, loss: 0.06571725010871887
step: 160, loss: 0.06278405338525772
step: 170, loss: 0.006753807421773672
step: 180, loss: 0.12606970965862274
step: 190, loss: 0.03840116038918495
step: 200, loss: 0.019184496253728867
step: 210, loss: 0.012241905555129051
/usr/local/lib/python3.7/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=0.5911708253358925, f1=0.6118546845124283,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03551917150616646
step: 10, loss: 0.04278747737407684
step: 20, loss: 0.014411550015211105
step: 30, loss: 0.001838164054788649
step: 40, loss: 0.06521841883659363
step: 50, loss: 0.10279514640569687
step: 60, loss: 0.14683392643928528
step: 70, loss: 0.01333583053201437
step: 80, loss: 0.08009949326515198
step: 90, loss: 0.010130398906767368
step: 100, loss: 0.09462790936231613
step: 110, loss: 0.04997800290584564
step: 120, loss: 0.019794659689068794
step: 130, loss: 0.006079118698835373
step: 140, loss: 0.07365964353084564
step: 150, loss: 0.13887859880924225
step: 160, loss: 0.04023422673344612
step: 170, loss: 0.04524162411689758
step: 180, loss: 0.05215667560696602
step: 190, loss: 0.021766401827335358
step: 200, loss: 0.00527178542688489
step: 210, loss: 0.1349562704563141
/usr/local/lib/python3.7/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=0.5833333333333333, f1=0.5838779956427015,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0400240458548069
step: 10, loss: 0.07427675276994705
step: 20, loss: 0.013974439352750778
step: 30, loss: 0.016434194520115852
step: 40, loss: 0.07276245206594467
step: 50, loss: 0.2664376497268677
step: 60, loss: 0.07128296047449112
step: 70, loss: 0.03620029613375664
step: 80, loss: 0.037850793451070786
step: 90, loss: 0.050959475338459015
step: 100, loss: 0.07600944489240646
step: 110, loss: 0.04962853342294693
step: 120, loss: 0.00402718223631382
step: 130, loss: 0.015062617138028145
step: 140, loss: 0.015807582065463066
step: 150, loss: 0.04623973369598389
step: 160, loss: 0.005137254949659109
step: 170, loss: 0.007451658137142658
step: 180, loss: 0.04835083708167076
step: 190, loss: 0.00463412469252944
step: 200, loss: 0.08840131759643555
step: 210, loss: 0.05357997119426727
/usr/local/lib/python3.7/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=0.576923076923077, f1=0.603415559772296,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02842884138226509
step: 10, loss: 0.054489586502313614
step: 20, loss: 0.0049280510284006596
step: 30, loss: 0.11490694433450699
step: 40, loss: 0.0054540918208658695
step: 50, loss: 0.007783061359077692
step: 60, loss: 0.0632421150803566
step: 70, loss: 0.039779532700777054
step: 80, loss: 0.08183122426271439
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.10375906527042389
step: 100, loss: 0.025289155542850494
step: 110, loss: 0.01363386120647192
step: 120, loss: 0.023928984999656677
step: 130, loss: 0.07000978291034698
step: 140, loss: 0.1137956753373146
step: 150, loss: 0.0553729347884655

```

```

step: 160, loss: 0.07058236747980118
step: 170, loss: 0.008585772477090359
step: 180, loss: 0.040805261582136154
step: 190, loss: 0.26701271533966064
step: 200, loss: 0.016408493742346764
step: 210, loss: 0.11495840549468994
/usr/local/lib/python3.7/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=0.583170254403131, f1=0.5877862595419848,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022121572867035866
step: 10, loss: 0.012404796667397022
step: 20, loss: 0.03696329519152641
step: 30, loss: 0.008150286041200161
step: 40, loss: 0.012056923471391201
step: 50, loss: 0.0024900222197175026
step: 60, loss: 0.07579828053712845
step: 70, loss: 0.008149194531142712
step: 80, loss: 0.10163899511098862
step: 90, loss: 0.014450409449636936
step: 100, loss: 0.05545773729681969
step: 110, loss: 0.043803371489048004
step: 120, loss: 0.0460045300424099
step: 130, loss: 0.012823894619941711
step: 140, loss: 0.013694755733013153
step: 150, loss: 0.0012600196059793234
step: 160, loss: 0.005689549725502729
step: 170, loss: 0.045832559466362
step: 180, loss: 0.005027105566114187
step: 190, loss: 0.012184408493340015
step: 200, loss: 0.0023869189899414778
step: 210, loss: 0.0023014270700514317
/usr/local/lib/python3.7/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=0.5708582834331337, f1=0.5776031434184675,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02461201697587967
step: 10, loss: 0.003672220977023244
step: 20, loss: 0.03645826503634453
step: 30, loss: 0.06274808943271637

```

```

step: 40, loss: 0.08967742323875427
step: 50, loss: 0.002330037532374263
step: 60, loss: 0.07918699085712433
step: 70, loss: 0.015210231766104698
step: 80, loss: 0.0033182334154844284
step: 90, loss: 0.025355264544487
step: 100, loss: 0.042204126715660095
step: 110, loss: 0.003922413568943739
step: 120, loss: 0.04869980737566948
step: 130, loss: 0.004273551050573587
step: 140, loss: 0.0012670510914176702
step: 150, loss: 0.08631602674722672
step: 160, loss: 0.020300183445215225
step: 170, loss: 0.010893880389630795
step: 180, loss: 0.016984963789582253
step: 190, loss: 0.016049016267061234
step: 200, loss: 0.012290786020457745
step: 210, loss: 0.0020288312807679176
/usr/local/lib/python3.7/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=0.5666666666666667, f1=0.5802469135802469,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013363828882575035
step: 10, loss: 0.00038452379521913826
step: 20, loss: 0.08373091369867325
step: 30, loss: 0.19294269382953644
step: 40, loss: 0.014437548816204071
step: 50, loss: 0.005449514836072922
step: 60, loss: 0.001352564780972898
step: 70, loss: 0.02738841623067856
step: 80, loss: 0.00727388309314847
step: 90, loss: 0.0005470588803291321
step: 100, loss: 0.002218283014371991
step: 110, loss: 0.012102843262255192
step: 120, loss: 0.049525875598192215
step: 130, loss: 0.0560002401471138
step: 140, loss: 0.013239783234894276
step: 150, loss: 0.017505327239632607
step: 160, loss: 0.02651493065059185
step: 170, loss: 0.07348839938640594
step: 180, loss: 0.18401531875133514
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.0053323195315897465
step: 200, loss: 0.002662640530616045

```

```

step: 210, loss: 0.005958002991974354
/usr/local/lib/python3.7/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=0.5625000000000001, f1=0.5627376425855514,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063949753530323505
step: 10, loss: 0.037266697734594345
step: 20, loss: 0.005555859301239252
step: 30, loss: 0.038072213530540466
step: 40, loss: 0.01370995119214058
step: 50, loss: 0.03592122346162796
step: 60, loss: 0.006484565325081348
step: 70, loss: 0.02848643809556961
step: 80, loss: 0.006734847091138363
step: 90, loss: 0.0069362446665763855
step: 100, loss: 0.03720000013709068
step: 110, loss: 0.11054122447967529
step: 120, loss: 0.007277941331267357
step: 130, loss: 0.005659548565745354
step: 140, loss: 0.04811783507466316
step: 150, loss: 0.013542412780225277
step: 160, loss: 0.01634407229721546
step: 170, loss: 0.018499698489904404
step: 180, loss: 0.0014622207963839173
step: 190, loss: 0.004960612393915653
step: 200, loss: 0.027959268540143967
step: 210, loss: 0.004700921941548586
/usr/local/lib/python3.7/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=0.5689655172413792, f1=0.5842217484008528,
best_f1=0.6150627615062761
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005664885509759188
step: 10, loss: 0.01443135179579258
step: 20, loss: 0.011330030858516693
step: 30, loss: 0.027116132900118828
step: 40, loss: 0.012379948981106281
step: 50, loss: 0.0017625009641051292
step: 60, loss: 0.053304076194763184
step: 70, loss: 0.005041378550231457
step: 80, loss: 0.0017813952872529626

```

```

step: 90, loss: 0.007368307560682297
step: 100, loss: 0.013345913961529732
step: 110, loss: 0.006330669391900301
step: 120, loss: 0.009726883843541145
step: 130, loss: 0.13836796581745148
step: 140, loss: 0.0014850231818854809
step: 150, loss: 0.0008521999116055667
step: 160, loss: 0.022389844059944153
step: 170, loss: 0.00043754972284659743
step: 180, loss: 0.0014113446231931448
step: 190, loss: 0.004579621367156506
step: 200, loss: 0.008729131892323494
step: 210, loss: 0.02431696653366089
/usr/local/lib/python3.7/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=0.5695652173913044, f1=0.5824411134903641,
best_f1=0.6150627615062761
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 [00:08, 279.33it/s]
load_f1 = 0.5984555984555985
real_f1 = 0.5984555984555985
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.46it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

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

Defaults for this optimization level are:
 enabled : True

```

opt_level          : 02
cast_model_type    : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights     : True
loss_scale         : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled           : True
opt_level         : 02
cast_model_type   : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32 : True
master_weights    : True
loss_scale        : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
    "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.5860747694969177
step: 10, loss: 0.3699457049369812
step: 20, loss: 0.29852885007858276
step: 30, loss: 0.43563222885131836
step: 40, loss: 0.4283563494682312
step: 50, loss: 0.28739842772483826
step: 60, loss: 0.26915502548217773
step: 70, loss: 0.2721223533153534
step: 80, loss: 0.3803044855594635
step: 90, loss: 0.2921803593635559
step: 100, loss: 0.3123225271701813
step: 110, loss: 0.37398913502693176
step: 120, loss: 0.07097073644399643
step: 130, loss: 0.16173309087753296
step: 140, loss: 0.12701542675495148
step: 150, loss: 0.1784985214471817
step: 160, loss: 0.12768356502056122

```

```

step: 170, loss: 0.21265456080436707
step: 180, loss: 0.037735968828201294
step: 190, loss: 0.19754242897033691
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.689655172413793, f1=0.7200000000000001,
best_f1=0.7200000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.41063982248306274
step: 10, loss: 0.13873662054538727
step: 20, loss: 0.14243528246879578
step: 30, loss: 0.09658415615558624
step: 40, loss: 0.09155689924955368
step: 50, loss: 0.1715438961982727
step: 60, loss: 0.2296360433101654
step: 70, loss: 0.18812036514282227
step: 80, loss: 0.11694139242172241
step: 90, loss: 0.174770787358284
step: 100, loss: 0.01404343731701374
step: 110, loss: 0.08454569429159164
step: 120, loss: 0.23567049205303192
step: 130, loss: 0.025750569999217987
step: 140, loss: 0.03462538868188858
step: 150, loss: 0.11624882370233536
step: 160, loss: 0.06506526470184326
step: 170, loss: 0.08335327357053757
step: 180, loss: 0.2164388746023178
step: 190, loss: 0.057422738522291183
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6983240223463687, f1=0.7138964577656676,
best_f1=0.7138964577656676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05833734944462776

```

```

step: 10, loss: 0.215668722987175
step: 20, loss: 0.018682807683944702
step: 30, loss: 0.07680007070302963
step: 40, loss: 0.031001750379800797
step: 50, loss: 0.03426777198910713
step: 60, loss: 0.030910154804587364
step: 70, loss: 0.20360040664672852
step: 80, loss: 0.0737101137638092
step: 90, loss: 0.06317804008722305
step: 100, loss: 0.08043084293603897
step: 110, loss: 0.00754470145329833
step: 120, loss: 0.009222608990967274
step: 130, loss: 0.004132918547838926
step: 140, loss: 0.060433946549892426
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.1384047120809555
step: 160, loss: 0.08051764219999313
step: 170, loss: 0.05051286518573761
step: 180, loss: 0.03823181614279747
step: 190, loss: 0.1117570698261261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7616580310880829, f1=0.7708333333333334,
best_f1=0.7708333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01942298375070095
step: 10, loss: 0.1487920582294464
step: 20, loss: 0.01561746746301651
step: 30, loss: 0.06942711025476456
step: 40, loss: 0.033753324300050735
step: 50, loss: 0.07632255554199219
step: 60, loss: 0.013407742604613304
step: 70, loss: 0.06238960102200508
step: 80, loss: 0.09247173368930817
step: 90, loss: 0.0668671727180481
step: 100, loss: 0.1416480988264084
step: 110, loss: 0.004992746748030186
step: 120, loss: 0.016437828540802002
step: 130, loss: 0.20305849611759186
step: 140, loss: 0.021932493895292282
step: 150, loss: 0.1256922334432602

```

```

step: 160, loss: 0.022877370938658714
step: 170, loss: 0.08009079098701477
step: 180, loss: 0.021593105047941208
step: 190, loss: 0.13653624057769775
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7819148936170213, f1=0.7431693989071039,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.046867743134498596
step: 10, loss: 0.008822151459753513
step: 20, loss: 0.062441933900117874
step: 30, loss: 0.00897202454507351
step: 40, loss: 0.1870037317276001
step: 50, loss: 0.040516551584005356
step: 60, loss: 0.20056934654712677
step: 70, loss: 0.028555743396282196
step: 80, loss: 0.014274933375418186
step: 90, loss: 0.007344079669564962
step: 100, loss: 0.003958349581807852
step: 110, loss: 0.012727775610983372
step: 120, loss: 0.027578184381127357
step: 130, loss: 0.030793873593211174
step: 140, loss: 0.03615182638168335
step: 150, loss: 0.018960123881697655
step: 160, loss: 0.003627296071499586
step: 170, loss: 0.0017280913889408112
step: 180, loss: 0.17828652262687683
step: 190, loss: 0.08885205537080765
/usr/local/lib/python3.7/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=0.7679558011049724, f1=0.7818696883852692,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006504178512841463
step: 10, loss: 0.02468474954366684
step: 20, loss: 0.009791286662220955
step: 30, loss: 0.0027160763274878263

```

```

step: 40, loss: 0.08111745864152908
step: 50, loss: 0.02369275689125061
step: 60, loss: 0.0021611680276691914
step: 70, loss: 0.009452786296606064
step: 80, loss: 0.03934541717171669
step: 90, loss: 0.004988270346075296
step: 100, loss: 0.0012097805738449097
step: 110, loss: 0.015886541455984116
step: 120, loss: 0.1272667497396469
step: 130, loss: 0.0294302087277174
step: 140, loss: 0.0959126204252243
step: 150, loss: 0.001304954057559371
step: 160, loss: 0.033749669790267944
step: 170, loss: 0.014543374069035053
step: 180, loss: 0.006596689578145742
step: 190, loss: 0.04346676915884018
/usr/local/lib/python3.7/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=0.7801047120418847, f1=0.779291553133515,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022686293814331293
step: 10, loss: 0.002287331037223339
step: 20, loss: 0.0457039438188076
step: 30, loss: 0.04504550248384476
step: 40, loss: 0.004516434855759144
step: 50, loss: 0.06984587758779526
step: 60, loss: 0.0845959410071373
step: 70, loss: 0.0014829599531367421
step: 80, loss: 0.006699014455080032
step: 90, loss: 0.01141088455915451
step: 100, loss: 0.0010171085596084595
step: 110, loss: 0.005318324547261
step: 120, loss: 0.0007899105548858643
step: 130, loss: 0.0012393315555527806
step: 140, loss: 0.0039890725165605545
step: 150, loss: 0.008620476350188255
step: 160, loss: 0.09691125154495239
step: 170, loss: 0.009142562747001648
step: 180, loss: 0.003379608504474163
step: 190, loss: 0.07783334702253342
/usr/local/lib/python3.7/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=0.7519582245430808, f1=0.7637362637362638,

```

```

best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003052895423024893
step: 10, loss: 0.0039247204549610615
step: 20, loss: 0.005507759749889374
step: 30, loss: 0.026896072551608086
step: 40, loss: 0.004243208095431328
step: 50, loss: 0.004308126866817474
step: 60, loss: 0.017608920112252235
step: 70, loss: 0.19182553887367249
step: 80, loss: 0.0008633402176201344
step: 90, loss: 0.003412820165976882
step: 100, loss: 0.049580566585063934
step: 110, loss: 0.0031211257446557283
step: 120, loss: 0.002461896976456046
step: 130, loss: 0.0028555605094879866
step: 140, loss: 0.005869029089808464
step: 150, loss: 0.025579817593097687
step: 160, loss: 0.005839771591126919
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.0064052981324493885
step: 180, loss: 0.11745210736989975
step: 190, loss: 0.0669386237859726
/usr/local/lib/python3.7/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=0.7601078167115903, f1=0.7555555555555554,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008186392486095428
step: 10, loss: 0.05752887576818466
step: 20, loss: 0.027306215837597847
step: 30, loss: 0.007338060066103935
step: 40, loss: 0.0050886753015220165
step: 50, loss: 0.0018842354184016585
step: 60, loss: 0.018373854458332062
step: 70, loss: 0.0009727111901156604
step: 80, loss: 0.005839458666741848
step: 90, loss: 0.005159558262676001
step: 100, loss: 0.02568749152123928
step: 110, loss: 0.10586817562580109
step: 120, loss: 0.0018366453005000949
step: 130, loss: 0.0029012365266680717
step: 140, loss: 0.022274741902947426

```

```

step: 150, loss: 0.16560621559619904
step: 160, loss: 0.0005013669142499566
step: 170, loss: 0.0015891434159129858
step: 180, loss: 0.011437413282692432
step: 190, loss: 0.001739352010190487
/usr/local/lib/python3.7/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=0.7631578947368421, f1=0.7431693989071039,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16636165976524353
step: 10, loss: 0.017492400482296944
step: 20, loss: 0.0657305121421814
step: 30, loss: 0.0016672409838065505
step: 40, loss: 0.07065758854150772
step: 50, loss: 0.052130743861198425
step: 60, loss: 0.0014792007859796286
step: 70, loss: 0.0009025081526488066
step: 80, loss: 0.013024871237576008
step: 90, loss: 0.0019767179619520903
step: 100, loss: 0.002733079483732581
step: 110, loss: 0.0037193787284195423
step: 120, loss: 0.11143873631954193
step: 130, loss: 0.0005871022003702819
step: 140, loss: 0.0009244380635209382
step: 150, loss: 0.0015112493420019746
step: 160, loss: 0.19680407643318176
step: 170, loss: 0.0008870505262166262
step: 180, loss: 0.0014651117380708456
step: 190, loss: 0.0013852189294993877
/usr/local/lib/python3.7/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=0.765498652291105, f1=0.7415730337078652,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002255016705021262
step: 10, loss: 0.003128194948658347
step: 20, loss: 0.29143571853637695
step: 30, loss: 0.002280828543007374
step: 40, loss: 0.002542814938351512
step: 50, loss: 0.0030016798991709948
step: 60, loss: 0.0022001040633767843

```



```

step: 70, loss: 0.002009608317166567
step: 80, loss: 0.001601871568709612
step: 90, loss: 0.0034694448113441467
step: 100, loss: 0.0004064622044097632
step: 110, loss: 0.0010587325086817145
step: 120, loss: 0.0019393155816942453
step: 130, loss: 0.04219285398721695
step: 140, loss: 0.002045291243121028
step: 150, loss: 0.0009356517693959177
step: 160, loss: 0.0017966885352507234
step: 170, loss: 0.0007658274844288826
step: 180, loss: 0.007486349903047085
step: 190, loss: 0.004081686958670616
/usr/local/lib/python3.7/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=0.7567567567567567, f1=0.752808988764045,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009429504163563251
step: 10, loss: 0.0026659017894417048
step: 20, loss: 0.0006783903809264302
step: 30, loss: 0.0009217248880304396
step: 40, loss: 0.0008160057477653027
step: 50, loss: 0.0019954906310886145
step: 60, loss: 0.0009332456975243986
step: 70, loss: 0.0031424607150256634
step: 80, loss: 0.003221542574465275
step: 90, loss: 0.004189217463135719
step: 100, loss: 0.0008922258857637644
step: 110, loss: 0.004527769051492214
step: 120, loss: 0.012177322059869766
step: 130, loss: 0.0210725087672472
step: 140, loss: 0.00783412903547287
step: 150, loss: 0.00268439669162035
step: 160, loss: 0.00044219029950909317
step: 170, loss: 0.0020610217470675707
step: 180, loss: 0.004302562214434147
step: 190, loss: 0.003243881743401289
/usr/local/lib/python3.7/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=0.7724867724867726, f1=0.7771739130434783,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017627837136387825
step: 10, loss: 0.003664088435471058
step: 20, loss: 0.0008972655050456524
step: 30, loss: 0.0059359995648264885
step: 40, loss: 0.0065207090228796005
step: 50, loss: 0.03792504966259003
step: 60, loss: 0.0008346408721990883
step: 70, loss: 0.00223959656432271
step: 80, loss: 0.001666333177126944
step: 90, loss: 0.008031199686229229
step: 100, loss: 0.029031898826360703
step: 110, loss: 0.004958544857800007
step: 120, loss: 0.0022442257031798363
step: 130, loss: 0.005948720965534449
step: 140, loss: 0.0004100005317013711
step: 150, loss: 0.0004884991212747991
step: 160, loss: 0.0018079282017424703
step: 170, loss: 0.00043439221917651594
step: 180, loss: 0.0006128231179900467
step: 190, loss: 0.029949422925710678
/usr/local/lib/python3.7/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=0.7624020887728459, f1=0.7472527472527473,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011455550789833069
step: 10, loss: 0.00039734956226311624
step: 20, loss: 0.0006368986214511096
step: 30, loss: 0.0006286200950853527
step: 40, loss: 0.0010392883559688926
step: 50, loss: 0.0009436543332412839
step: 60, loss: 0.12271852046251297
step: 70, loss: 0.01099249254912138
step: 80, loss: 0.0026731824036687613
step: 90, loss: 0.003856358351185918
step: 100, loss: 0.021517563611268997
step: 110, loss: 0.0017451152671128511
step: 120, loss: 0.00627576420083642
step: 130, loss: 0.0011697817826643586
step: 140, loss: 0.0008434601477347314
step: 150, loss: 0.0007123261457309127
step: 160, loss: 0.0031684606801718473
step: 170, loss: 0.0009381055715493858
step: 180, loss: 0.005667706951498985

```

```

step: 190, loss: 0.0010379317682236433
/usr/local/lib/python3.7/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=0.7567567567567567, f1=0.7605633802816902,
best_f1=0.7431693989071039
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008934746496379375
step: 10, loss: 0.0006673657917417586
step: 20, loss: 0.05470908433198929
step: 30, loss: 0.0013476377353072166
step: 40, loss: 0.0011005504056811333
step: 50, loss: 0.00351284584030509
step: 60, loss: 0.002062340034171939
step: 70, loss: 0.0008644303889013827
step: 80, loss: 0.0026942032855004072
step: 90, loss: 0.001687712618149817
step: 100, loss: 0.001342430361546576
step: 110, loss: 0.001384460600093007
step: 120, loss: 0.0005211810348555446
step: 130, loss: 0.00021263725648168474
step: 140, loss: 0.005351543892174959
step: 150, loss: 0.00047860541963018477
step: 160, loss: 0.005418866407126188
step: 170, loss: 0.001593775232322514
step: 180, loss: 0.015473474748432636
step: 190, loss: 0.0003995979204773903
/usr/local/lib/python3.7/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=0.7573333333333333, f1=0.7423822714681441,
best_f1=0.7431693989071039
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 170.28it/s]
load_f1 = 0.6574585635359116
real_f1 = 0.646900269541779
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 193.25it/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 bert-base-uncased \
--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 bert-base-uncased \
  --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: 0.6208709478378296
step: 10, loss: 0.3753649890422821
step: 20, loss: 0.3145114779472351
step: 30, loss: 0.37894412875175476
step: 40, loss: 0.2788117825984955
step: 50, loss: 0.2894257605075836
step: 60, loss: 0.2890865206718445
step: 70, loss: 0.36827483773231506
step: 80, loss: 0.35561051964759827
step: 90, loss: 0.2311665117740631
step: 100, loss: 0.22713589668273926
step: 110, loss: 0.290507048368454
step: 120, loss: 0.20087595283985138
step: 130, loss: 0.023152196779847145
step: 140, loss: 0.1857108175754547
step: 150, loss: 0.2593490779399872
step: 160, loss: 0.10743484646081924
step: 170, loss: 0.26644158363342285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, f1=0.6965174129353233,
best_f1=0.6965174129353233
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11205919086933136
step: 10, loss: 0.23630967736244202
step: 20, loss: 0.08145750313997269
step: 30, loss: 0.13673678040504456
step: 40, loss: 0.08426313102245331
step: 50, loss: 0.10212019830942154
step: 60, loss: 0.10292063653469086
step: 70, loss: 0.09505636990070343
step: 80, loss: 0.11131690442562103
step: 90, loss: 0.11657392978668213
step: 100, loss: 0.14209921658039093
step: 110, loss: 0.029399225488305092
step: 120, loss: 0.04871276393532753
step: 130, loss: 0.08565958589315414
step: 140, loss: 0.3038383424282074
step: 150, loss: 0.15978612005710602

```

```

step: 160, loss: 0.14384225010871887
step: 170, loss: 0.0740838423371315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7676767676767677, f1=0.7628361858190709,
best_f1=0.7628361858190709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04589603841304779
step: 10, loss: 0.027856187894940376
step: 20, loss: 0.13310706615447998
step: 30, loss: 0.22038309276103973
step: 40, loss: 0.14459441602230072
step: 50, loss: 0.07948023825883865
step: 60, loss: 0.12298727035522461
step: 70, loss: 0.06199325621128082
step: 80, loss: 0.07122794538736343
step: 90, loss: 0.18245084583759308
step: 100, loss: 0.021113421767950058
step: 110, loss: 0.07818453758955002
step: 120, loss: 0.054652128368616104
step: 130, loss: 0.04032561182975769
step: 140, loss: 0.14661964774131775
step: 150, loss: 0.11500699073076248
step: 160, loss: 0.07093897461891174
step: 170, loss: 0.14762355387210846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7786666666666666, f1=0.772378516624041,
best_f1=0.772378516624041
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019822267815470695
step: 10, loss: 0.027099167928099632
step: 20, loss: 0.018666738644242287
step: 30, loss: 0.05891574174165726

```

```

step: 40, loss: 0.00870172493159771
step: 50, loss: 0.01765969954431057
step: 60, loss: 0.2933269143104553
step: 70, loss: 0.014387513510882854
step: 80, loss: 0.07886816561222076
step: 90, loss: 0.030948594212532043
step: 100, loss: 0.1963164508342743
step: 110, loss: 0.04092586785554886
step: 120, loss: 0.01880648173391819
step: 130, loss: 0.029308408498764038
step: 140, loss: 0.03623213618993759
step: 150, loss: 0.03920881822705269
step: 160, loss: 0.03964788094162941
step: 170, loss: 0.0043827807530760765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.824742268041237, f1=0.7938931297709925,
best_f1=0.7938931297709925
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03050174191594124
step: 10, loss: 0.0404733307659626
step: 20, loss: 0.01805218495428562
step: 30, loss: 0.03225826099514961
step: 40, loss: 0.01861579157412052
step: 50, loss: 0.09967617690563202
step: 60, loss: 0.029553476721048355
step: 70, loss: 0.3470028340816498
step: 80, loss: 0.0931498259305954
step: 90, loss: 0.13062599301338196
step: 100, loss: 0.08470024913549423
step: 110, loss: 0.2846479117870331
step: 120, loss: 0.00516888964921236
step: 130, loss: 0.11958368867635727
step: 140, loss: 0.033746588975191116
step: 150, loss: 0.07854551821947098
step: 160, loss: 0.030839405953884125
step: 170, loss: 0.006890848278999329
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```


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=0.8303797468354431, f1=0.7651331719128328,
best_f1=0.7651331719128328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003591670421883464
step: 10, loss: 0.00746597396209836
step: 20, loss: 0.030872199684381485
step: 30, loss: 0.0010925120441243052
step: 40, loss: 0.048856835812330246
step: 50, loss: 0.11108041554689407
step: 60, loss: 0.07734636962413788
step: 70, loss: 0.09026351571083069
step: 80, loss: 0.022172143682837486
step: 90, loss: 0.035639479756355286
step: 100, loss: 0.006063130684196949
step: 110, loss: 0.0019254995277151465
step: 120, loss: 0.003400791436433792
step: 130, loss: 0.01103877928107977
step: 140, loss: 0.006603463087230921
step: 150, loss: 0.05243257060647011
step: 160, loss: 0.047264859080314636
step: 170, loss: 0.0027761098463088274
/usr/local/lib/python3.7/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=0.8132992327365729, f1=0.7857142857142857,
best_f1=0.7651331719128328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015197602333500981
step: 10, loss: 0.012026231735944748
step: 20, loss: 0.008188840001821518
step: 30, loss: 0.013475471176207066
step: 40, loss: 0.010450302623212337
step: 50, loss: 0.07607221603393555
step: 60, loss: 0.0015976082067936659
step: 70, loss: 0.017284374684095383
step: 80, loss: 0.0439736470580101
step: 90, loss: 0.000997377559542656
step: 100, loss: 0.002947795670479536
step: 110, loss: 0.025622103363275528
step: 120, loss: 0.002751907566562295
step: 130, loss: 0.1393931359052658
```

```

step: 140, loss: 0.0078069837763905525
step: 150, loss: 0.011723031289875507
step: 160, loss: 0.005478072911500931
step: 170, loss: 0.03501461446285248
/usr/local/lib/python3.7/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=0.8, f1=0.7715736040609138, best_f1=0.7651331719128328
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01175597496330738
step: 10, loss: 0.005597907584160566
step: 20, loss: 0.00037359673297032714
step: 30, loss: 0.00561732379719615
step: 40, loss: 0.000769125996157527
step: 50, loss: 0.00405000289902091
step: 60, loss: 0.00143570511136204
step: 70, loss: 0.0008483501151204109
step: 80, loss: 0.001615081331692636
step: 90, loss: 0.006854633800685406
step: 100, loss: 0.029842497780919075
step: 110, loss: 0.09849821776151657
step: 120, loss: 0.0004921084619127214
step: 130, loss: 0.012810804881155491
step: 140, loss: 0.007701690308749676
step: 150, loss: 0.14436163008213043
step: 160, loss: 0.019262254238128662
step: 170, loss: 0.002525411080569029
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8426395939086295, f1=0.7909319899244334,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003716642677318305
step: 10, loss: 0.0011317559983581305
step: 20, loss: 0.002686471212655306
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.0023182793520390987
step: 40, loss: 0.0005480466643348336
step: 50, loss: 0.00043408581404946744

```

```

step: 60, loss: 0.01959356479346752
step: 70, loss: 0.00823761522769928
step: 80, loss: 0.0004151341854594648
step: 90, loss: 0.01595686934888363
step: 100, loss: 0.0032221514265984297
step: 110, loss: 0.0013594586635008454
step: 120, loss: 0.012378249317407608
step: 130, loss: 0.09747273474931717
step: 140, loss: 0.008356461301445961
step: 150, loss: 0.015401050448417664
step: 160, loss: 0.03812716156244278
step: 170, loss: 0.0005294220754876733
/usr/local/lib/python3.7/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=0.8404255319148938, f1=0.806282722513089,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021366257220506668
step: 10, loss: 0.0002270653931191191
step: 20, loss: 0.18337631225585938
step: 30, loss: 0.0004061675281263888
step: 40, loss: 0.0003267905267421156
step: 50, loss: 0.0815521627664566
step: 60, loss: 0.00017530554032418877
step: 70, loss: 0.009048041887581348
step: 80, loss: 0.0021777863148599863
step: 90, loss: 0.013215244747698307
step: 100, loss: 0.0004860087647102773
step: 110, loss: 0.0026094221975654364
step: 120, loss: 0.0006409049965441227
step: 130, loss: 0.0013032922288402915
step: 140, loss: 0.13275951147079468
step: 150, loss: 0.04247702658176422
step: 160, loss: 0.02069278247654438
step: 170, loss: 0.00020562621648423374
/usr/local/lib/python3.7/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=0.8232445520581113, f1=0.7846889952153109,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047709252685308456
step: 10, loss: 0.02832961082458496

```

```

step: 20, loss: 0.001509928610175848
step: 30, loss: 0.0001845131191657856
step: 40, loss: 0.0007578147924505174
step: 50, loss: 0.0024931272491812706
step: 60, loss: 0.03443383425474167
step: 70, loss: 0.0018471510848030448
step: 80, loss: 0.00034488795790821314
step: 90, loss: 0.0019309193594381213
step: 100, loss: 0.012068388052284718
step: 110, loss: 0.008913157507777214
step: 120, loss: 0.0009526594076305628
step: 130, loss: 0.00025050260592252016
step: 140, loss: 0.04953882843255997
step: 150, loss: 0.03150627762079239
step: 160, loss: 0.05462475121021271
step: 170, loss: 0.0005708563257940114
/usr/local/lib/python3.7/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=0.8238341968911919, f1=0.7786259541984734,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01417714636772871
step: 10, loss: 0.0005116310785524547
step: 20, loss: 0.0003728649753611535
step: 30, loss: 0.0008686677319929004
step: 40, loss: 0.00019269609765615314
step: 50, loss: 0.0002417021314613521
step: 60, loss: 0.00024845157167874277
step: 70, loss: 0.06670600920915604
step: 80, loss: 8.931061893235892e-05
step: 90, loss: 0.0005216005374677479
step: 100, loss: 0.0010434574214741588
step: 110, loss: 0.0002787198463920504
step: 120, loss: 0.009932305663824081
step: 130, loss: 0.0038695475086569786
step: 140, loss: 0.0045863487757742405
step: 150, loss: 0.007277729921042919
step: 160, loss: 0.00016721864813007414
step: 170, loss: 0.0003010371292475611
/usr/local/lib/python3.7/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=0.8125, f1=0.7575757575757577, best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00549383694306016
step: 10, loss: 0.00047069863649085164
step: 20, loss: 0.00032019056379795074
step: 30, loss: 0.00010373349505243823
step: 40, loss: 0.0007931647705845535
step: 50, loss: 0.0002398498763795942
step: 60, loss: 0.0008482017437927425
step: 70, loss: 0.0027497424744069576
step: 80, loss: 0.0012505070772022009
step: 90, loss: 0.00025433924747630954
step: 100, loss: 0.0007609547465108335
step: 110, loss: 0.00019516705651767552
step: 120, loss: 0.031779058277606964
step: 130, loss: 0.0001375235733576119
step: 140, loss: 0.0002653195697348565
step: 150, loss: 0.0002600336738396436
step: 160, loss: 0.00017939312965609133
step: 170, loss: 0.0006164719234220684
/usr/local/lib/python3.7/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=0.8043478260869564, f1=0.7745358090185676,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007303419406525791
step: 10, loss: 0.00012085065827704966
step: 20, loss: 0.0005406226846389472
step: 30, loss: 0.00018069315410684794
step: 40, loss: 0.0001838218013290316
step: 50, loss: 0.00022516574244946241
step: 60, loss: 9.111643157666549e-05
step: 70, loss: 0.00117797392886132
step: 80, loss: 0.007795966230332851
step: 90, loss: 0.00022445245122071356
step: 100, loss: 0.03776091709733009
step: 110, loss: 0.0005791764706373215
step: 120, loss: 0.0280525591224432
step: 130, loss: 0.0015535939019173384
step: 140, loss: 0.020424626767635345
step: 150, loss: 0.00025025333161465824
step: 160, loss: 0.00010495173773961142
step: 170, loss: 0.00037377458647824824
/usr/local/lib/python3.7/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=0.8132992327365729, f1=0.7676767676767677,
best_f1=0.7909319899244334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007463428191840649
step: 10, loss: 0.004169286694377661
step: 20, loss: 0.06057625263929367
step: 30, loss: 0.0012093756813555956
step: 40, loss: 0.0004732007801067084
step: 50, loss: 0.00016581623640377074
step: 60, loss: 0.017158757895231247
step: 70, loss: 0.0003416991967242211
step: 80, loss: 0.02809653989970684
step: 90, loss: 0.0009789559990167618
step: 100, loss: 0.0015559608582407236
step: 110, loss: 0.0002416410279693082
step: 120, loss: 0.0006464757025241852
step: 130, loss: 0.0040021250024437904
step: 140, loss: 0.00036808487493544817
step: 150, loss: 0.0005902992561459541
step: 160, loss: 0.00016263349971268326
step: 170, loss: 0.00028634440968744457
/usr/local/lib/python3.7/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=0.8142493638676845, f1=0.7611940298507462,
best_f1=0.7909319899244334
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:09, 202.25it/s]
load_f1 = 0.5353846153846153
real_f1 = 0.524390243902439
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, 184.26it/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 bert-base-uncased \
    --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 bert-base-uncased \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Downloading: 100% 433/433 [00:00<00:00, 352kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 255kB/s]
 Downloading: 100% 440M/440M [00:10<00:00, 41.7MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

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

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

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

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

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

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

step: 0, loss: 0.6313666105270386


```

step: 10, loss: 0.6160402894020081
step: 20, loss: 0.4453106224536896
step: 30, loss: 0.18677982687950134
step: 40, loss: 0.1482085883617401
step: 50, loss: 0.07969135046005249
step: 60, loss: 0.11881394684314728
step: 70, loss: 0.13898612558841705
step: 80, loss: 0.037479083985090256
step: 90, loss: 0.1138644739985466
step: 100, loss: 0.03148370236158371
step: 110, loss: 0.18885385990142822
step: 120, loss: 0.009103581309318542
step: 130, loss: 0.007923022843897343
step: 140, loss: 0.0045296926982700825
step: 150, loss: 0.01730893738567829
step: 160, loss: 0.013848387636244297
step: 170, loss: 0.06609407812356949
step: 180, loss: 0.05616101250052452
step: 190, loss: 0.09819313883781433
step: 200, loss: 0.1145409643650055
step: 210, loss: 0.007917859591543674
step: 220, loss: 0.0559203214943409
step: 230, loss: 0.020429883152246475
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9665924276169264, f1=0.9718785151856018,
best_f1=0.9718785151856018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009322626516222954
step: 10, loss: 0.016781968995928764
step: 20, loss: 0.048843588680028915
step: 30, loss: 0.2023492455482483
step: 40, loss: 0.08794018626213074
step: 50, loss: 0.035454485565423965
step: 60, loss: 0.007601082790642977
step: 70, loss: 0.024727115407586098
step: 80, loss: 0.0077914223074913025
step: 90, loss: 0.1364997923374176
step: 100, loss: 0.049655769020318985
step: 110, loss: 0.16063852608203888
step: 120, loss: 0.038879454135894775

```

```

step: 130, loss: 0.05648957937955856
step: 140, loss: 0.001224212464876473
step: 150, loss: 0.020025847479701042
step: 160, loss: 0.026653872802853584
step: 170, loss: 0.00235636904835701
step: 180, loss: 0.004518745467066765
step: 190, loss: 0.0032607463654130697
step: 200, loss: 0.002789855469018221
step: 210, loss: 0.03461331129074097
step: 220, loss: 0.12110108882188797
step: 230, loss: 0.01125344354659319
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9753363228699552, f1=0.976271186440678,
best_f1=0.976271186440678
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007171829231083393
step: 10, loss: 0.006613289937376976
step: 20, loss: 0.12185335159301758
step: 30, loss: 0.023588968440890312
step: 40, loss: 0.16844624280929565
step: 50, loss: 0.0038270633667707443
step: 60, loss: 0.0055871088989079
step: 70, loss: 0.005180764943361282
step: 80, loss: 0.0007737440173514187
step: 90, loss: 0.0036055578384548426
step: 100, loss: 0.0006038368446752429
step: 110, loss: 0.0009110231185331941
step: 120, loss: 0.003765536705031991
step: 130, loss: 0.00036347273271530867
step: 140, loss: 0.0009221050422638655
step: 150, loss: 0.009423513896763325
step: 160, loss: 0.10802282392978668
step: 170, loss: 0.009301080368459225
step: 180, loss: 0.006033124402165413
step: 190, loss: 0.004456007853150368
step: 200, loss: 0.08064481616020203
step: 210, loss: 0.010812790133059025
step: 220, loss: 0.0005625163903459907
step: 230, loss: 0.000646439497359097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776785714285714, f1=0.9730941704035874,
best_f1=0.9730941704035874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017781708447728306
step: 10, loss: 0.001118619809858501
step: 20, loss: 0.006761047523468733
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.0020740488544106483
step: 40, loss: 0.004742722027003765
step: 50, loss: 0.0003202592197339982
step: 60, loss: 0.003143803682178259
step: 70, loss: 0.000878685328643769
step: 80, loss: 0.0008694067364558578
step: 90, loss: 0.0021464533638209105
step: 100, loss: 0.0018053387757390738
step: 110, loss: 0.001844258513301611
step: 120, loss: 0.00542720640078187
step: 130, loss: 0.005393777042627335
step: 140, loss: 0.0033724645618349314
step: 150, loss: 0.17186100780963898
step: 160, loss: 0.004763667471706867
step: 170, loss: 0.003150732023641467
step: 180, loss: 0.0013604884734377265
step: 190, loss: 0.00291770719923079
step: 200, loss: 0.0016682051355019212
step: 210, loss: 0.0515204481780529
step: 220, loss: 0.000471106031909585
step: 230, loss: 0.003650605445727706
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9788182831661093,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0009844383457675576
step: 10, loss: 0.008379202336072922
step: 20, loss: 0.011639807373285294
step: 30, loss: 0.0002290497359354049
step: 40, loss: 0.0003848338674288243
step: 50, loss: 0.0008529890328645706
step: 60, loss: 0.20988187193870544
step: 70, loss: 0.00041096104541793466
step: 80, loss: 0.0007606280851177871
step: 90, loss: 0.001187873538583517
step: 100, loss: 0.0007678103866055608
step: 110, loss: 0.0005572218797169626
step: 120, loss: 0.00012593316205311567
step: 130, loss: 0.004529277794063091
step: 140, loss: 0.00030592118855565786
step: 150, loss: 0.00028016246506012976
step: 160, loss: 0.000293567223707214
step: 170, loss: 0.019679438322782516
step: 180, loss: 0.0011461434187367558
step: 190, loss: 0.018029630184173584
step: 200, loss: 0.006423930637538433
step: 210, loss: 0.0007857974851503968
step: 220, loss: 0.001188797876238823
step: 230, loss: 0.0013285784516483545
/usr/local/lib/python3.7/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=0.9733924611973392, f1=0.9720044792833147,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062195416539907455
step: 10, loss: 0.0007956397021189332
step: 20, loss: 0.004060571547597647
step: 30, loss: 0.00033373889164067805
step: 40, loss: 0.0002919159014709294
step: 50, loss: 0.00031014109845273197
step: 60, loss: 0.0002920467231888324
step: 70, loss: 0.003011043183505535
step: 80, loss: 0.0048033567145466805
step: 90, loss: 0.0005145580507814884
step: 100, loss: 0.06768209487199783
step: 110, loss: 0.004651864059269428
step: 120, loss: 0.0019059175392612815
step: 130, loss: 0.0006140762707218528
step: 140, loss: 0.0026206846814602613
step: 150, loss: 0.00169295200612396

```

```

step: 160, loss: 0.002045018831267953
step: 170, loss: 0.00044337729923427105
step: 180, loss: 0.07206191122531891
step: 190, loss: 0.07733612507581711
step: 200, loss: 0.01807139627635479
step: 210, loss: 0.0010269093327224255
step: 220, loss: 0.002383536659181118
step: 230, loss: 0.090727299451828
/usr/local/lib/python3.7/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=0.9707207207207207, f1=0.9660633484162895,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0048971399664878845
step: 10, loss: 0.00024032837245613337
step: 20, loss: 0.0021800464019179344
step: 30, loss: 0.016430726274847984
step: 40, loss: 0.0013684442965313792
step: 50, loss: 0.00041244574822485447
step: 60, loss: 0.002898563165217638
step: 70, loss: 0.017623094841837883
step: 80, loss: 0.00021868146723136306
step: 90, loss: 0.0002984184247907251
step: 100, loss: 0.00040385080501437187
step: 110, loss: 0.00044822331983596087
step: 120, loss: 0.00019920409249607474
step: 130, loss: 0.00043693528277799487
step: 140, loss: 0.0004216132510919124
step: 150, loss: 0.007178385276347399
step: 160, loss: 0.048633139580488205
step: 170, loss: 0.0018870962085202336
step: 180, loss: 0.04201589524745941
step: 190, loss: 0.03671328350901604
step: 200, loss: 0.035639889538288116
step: 210, loss: 0.00023776183661539108
step: 220, loss: 0.011402975767850876
step: 230, loss: 0.0006328829913400114
/usr/local/lib/python3.7/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=0.9820627802690582, f1=0.9797752808988766,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0005014449125155807
step: 10, loss: 0.0011985261226072907
step: 20, loss: 0.0015206370735540986
step: 30, loss: 0.00042068917537108064
step: 40, loss: 0.0028364837635308504
step: 50, loss: 0.0003247592831030488
step: 60, loss: 0.0010255769593641162
step: 70, loss: 0.0005445001879706979
step: 80, loss: 0.0007236028322950006
step: 90, loss: 0.00011035819625249133
step: 100, loss: 0.0004248509940225631
step: 110, loss: 0.00024639375624246895
step: 120, loss: 0.0027348354924470186
step: 130, loss: 0.0007153471233323216
step: 140, loss: 0.00012782312114723027
step: 150, loss: 0.00026213799719698727
step: 160, loss: 0.00044960726518183947
step: 170, loss: 0.00013201228284742683
step: 180, loss: 0.04124615341424942
step: 190, loss: 0.0007343311444856226
step: 200, loss: 0.0017126353923231363
step: 210, loss: 0.0005677210283465683
step: 220, loss: 0.00020683559705503285
step: 230, loss: 0.0007354849367402494
/usr/local/lib/python3.7/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=0.9776286353467561, f1=0.9753363228699552,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030860621482133865
step: 10, loss: 0.0012312246253713965
step: 20, loss: 0.00024086461053229868
step: 30, loss: 0.0001487628760514781
step: 40, loss: 0.015636755153536797
step: 50, loss: 7.094251486705616e-05
step: 60, loss: 0.00023654797405470163
step: 70, loss: 8.370511932298541e-05
step: 80, loss: 6.808075704611838e-05
step: 90, loss: 0.002329841023311019
step: 100, loss: 0.000330111215589568
step: 110, loss: 0.0031149580609053373
step: 120, loss: 4.049178824061528e-05
step: 130, loss: 9.718483488541096e-05
step: 140, loss: 0.006441451609134674
step: 150, loss: 0.00047514791367575526

```

```

step: 160, loss: 0.0014697522856295109
step: 170, loss: 0.023571917787194252
step: 180, loss: 0.0015559792518615723
step: 190, loss: 0.00019159408111590892
step: 200, loss: 0.00023734445858281106
step: 210, loss: 0.0002505175652913749
step: 220, loss: 7.328364154091105e-05
step: 230, loss: 0.00017466458666604012
/usr/local/lib/python3.7/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=0.9820224719101124, f1=0.9763779527559054,
best_f1=0.9788182831661093
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021487149933818728
step: 10, loss: 0.0003837365366052836
step: 20, loss: 0.0002708900428842753
step: 30, loss: 0.00013476524327415973
step: 40, loss: 5.68471041333396e-05
step: 50, loss: 0.00018210116832051426
step: 60, loss: 0.0017164844321087003
step: 70, loss: 0.00020430804579518735
step: 80, loss: 0.0001601203257450834
step: 90, loss: 0.00021539928275160491
step: 100, loss: 7.372339314315468e-05
step: 110, loss: 0.00024919735733419657
step: 120, loss: 0.0004546447889879346
step: 130, loss: 0.00010388806549599394
step: 140, loss: 0.02138170599937439
step: 150, loss: 0.04196241497993469
step: 160, loss: 0.00017778182518668473
step: 170, loss: 3.992929487139918e-05
step: 180, loss: 0.00031823391327634454
step: 190, loss: 0.04870704188942909
step: 200, loss: 5.982339644106105e-05
step: 210, loss: 0.0003956783621106297
step: 220, loss: 0.00030150319798849523
step: 230, loss: 0.00013670866610482335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9842696629213483, f1=0.9784335981838819,

```

```

best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018043439195025712
step: 10, loss: 0.00011102335702162236
step: 20, loss: 0.00013091655273456126
step: 30, loss: 0.0026319127064198256
step: 40, loss: 0.00014548587205354124
step: 50, loss: 0.0001791502145351842
step: 60, loss: 0.08259917050600052
step: 70, loss: 0.0002777956542558968
step: 80, loss: 6.303645932348445e-05
step: 90, loss: 0.0002525286690797657
step: 100, loss: 0.00016962597146630287
step: 110, loss: 0.00013901609054300934
step: 120, loss: 8.846498531056568e-05
step: 130, loss: 0.00020769701222889125
step: 140, loss: 4.764707773574628e-05
step: 150, loss: 0.001293735345825553
step: 160, loss: 8.999722194857895e-05
step: 170, loss: 0.007131515070796013
step: 180, loss: 0.001636813278310001
step: 190, loss: 4.9789712647907436e-05
step: 200, loss: 0.00016186882567126304
step: 210, loss: 8.151022484526038e-05
step: 220, loss: 6.385167216649279e-05
step: 230, loss: 0.00011343829100951552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9776286353467561, f1=0.9730941704035874,
best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.053914254764095e-05
step: 10, loss: 0.004032640717923641
step: 20, loss: 4.498757334658876e-05
step: 30, loss: 0.0001996853679884225
step: 40, loss: 0.0002986134495586157
step: 50, loss: 0.01539567019790411
step: 60, loss: 0.0008547070901840925
step: 70, loss: 6.536010914715007e-05
step: 80, loss: 0.0005130688077770174
step: 90, loss: 8.371299918508157e-05
step: 100, loss: 3.855476461467333e-05
step: 110, loss: 5.3725147154182196e-05

```



```

step: 120, loss: 7.248753536259755e-05
step: 130, loss: 4.370680107967928e-05
step: 140, loss: 6.02427389821969e-05
step: 150, loss: 0.00027363255503587425
step: 160, loss: 0.0012194202281534672
step: 170, loss: 5.398325447458774e-05
step: 180, loss: 0.0017191523220390081
step: 190, loss: 0.0001020272247842513
step: 200, loss: 7.554941112175584e-05
step: 210, loss: 5.214969860389829e-05
step: 220, loss: 3.750035830307752e-05
step: 230, loss: 0.0031827285420149565
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9809203142536477, f1=0.9751131221719457,
best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.69718356220983e-05
step: 10, loss: 7.471938442904502e-05
step: 20, loss: 0.00018874014494940639
step: 30, loss: 7.768873911118135e-05
step: 40, loss: 4.623085987986997e-05
step: 50, loss: 0.00017239006410818547
step: 60, loss: 0.0211234912276268
step: 70, loss: 2.997619958478026e-05
step: 80, loss: 6.284801929723471e-05
step: 90, loss: 0.0414714515209198
step: 100, loss: 4.1761930333450437e-05
step: 110, loss: 0.01311770360916853
step: 120, loss: 0.0009371191845275462
step: 130, loss: 4.947231718688272e-05
step: 140, loss: 5.723924550693482e-05
step: 150, loss: 5.993645027047023e-05
step: 160, loss: 0.00011904364509973675
step: 170, loss: 5.34715800313279e-05
step: 180, loss: 8.748114487389103e-05
step: 190, loss: 3.803321305895224e-05
step: 200, loss: 0.001654560910537839
step: 210, loss: 2.1013880541431718e-05
step: 220, loss: 0.00037385107134468853
step: 230, loss: 2.910474722739309e-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=0.9798206278026906, f1=0.9753363228699552,

```

```

best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003558448515832424
step: 10, loss: 0.00014171349175740033
step: 20, loss: 4.973264003638178e-05
step: 30, loss: 4.6481607569148764e-05
step: 40, loss: 0.022921158000826836
step: 50, loss: 9.960120951291174e-05
step: 60, loss: 3.3254393201787025e-05
step: 70, loss: 8.044371497817338e-05
step: 80, loss: 8.741364581510425e-05
step: 90, loss: 0.00042992105591110885
step: 100, loss: 4.969057044945657e-05
step: 110, loss: 0.0001878091279650107
step: 120, loss: 2.6377760150353424e-05
step: 130, loss: 3.5023927921429276e-05
step: 140, loss: 4.308854477130808e-05
step: 150, loss: 5.774123565061018e-05
step: 160, loss: 0.0007394130807369947
step: 170, loss: 4.17032242694404e-05
step: 180, loss: 0.00024534022668376565
step: 190, loss: 5.474625140777789e-05
step: 200, loss: 8.318566688103601e-05
step: 210, loss: 0.00011000559607055038
step: 220, loss: 3.9179562008939683e-05
step: 230, loss: 7.62989729992114e-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=0.9820627802690582, f1=0.9774774774774775,
best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7815929570351727e-05
step: 10, loss: 2.143846904800739e-05
step: 20, loss: 3.95906281482894e-05
step: 30, loss: 6.142690108390525e-05
step: 40, loss: 0.0001945612020790577
step: 50, loss: 0.0036008029710501432
step: 60, loss: 5.645798592013307e-05
step: 70, loss: 9.488880459684879e-05
step: 80, loss: 0.00012688004062511027
step: 90, loss: 9.682605741545558e-05
step: 100, loss: 0.00013741773727815598
step: 110, loss: 3.132472193101421e-05

```

```

step: 120, loss: 5.070889164926484e-05
step: 130, loss: 7.779243605909869e-05
step: 140, loss: 7.175366044975817e-05
step: 150, loss: 0.0001907076803036034
step: 160, loss: 3.701768218888901e-05
step: 170, loss: 2.72604502242757e-05
step: 180, loss: 0.0001508156128693372
step: 190, loss: 0.0001229803019668907
step: 200, loss: 0.000103120764833875
step: 210, loss: 0.0006455379771068692
step: 220, loss: 0.0006382195861078799
step: 230, loss: 4.207074016449042e-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=0.9820627802690582, f1=0.9774774774774775,
best_f1=0.9784335981838819
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 [00:14, 165.11it/s]

```

```
load_f1 = 0.9831271091113611
real_f1 = 0.9808773903262092
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, 184.44it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.636089026927948
step: 10, loss: 0.564772367477417
step: 20, loss: 0.4878000020980835
step: 30, loss: 0.17334645986557007
step: 40, loss: 0.21001403033733368
step: 50, loss: 0.25118738412857056
step: 60, loss: 0.103816919028759
step: 70, loss: 0.21679669618606567
step: 80, loss: 0.09102486073970795
step: 90, loss: 0.36530813574790955
step: 100, loss: 0.07199956476688385
step: 110, loss: 0.1492275446653366
step: 120, loss: 0.17446886003017426
step: 130, loss: 0.117298923432827
step: 140, loss: 0.11727374792098999
step: 150, loss: 0.060932599008083344
step: 160, loss: 0.029064994305372238
step: 170, loss: 0.24856676161289215

```

```

step: 180, loss: 0.17126227915287018
step: 190, loss: 0.021905740723013878
step: 200, loss: 0.15603764355182648
step: 210, loss: 0.02232777327299118
step: 220, loss: 0.27868038415908813
step: 230, loss: 0.14336155354976654
step: 240, loss: 0.019814414903521538
step: 250, loss: 0.009108795784413815
step: 260, loss: 0.07697809487581253
step: 270, loss: 0.0073485528118908405
step: 280, loss: 0.017690328881144524
step: 290, loss: 0.16017784178256989
step: 300, loss: 0.07453105598688126
step: 310, loss: 0.107426218688488
step: 320, loss: 0.13322660326957703
step: 330, loss: 0.03586847707629204
step: 340, loss: 0.026646055281162262
step: 350, loss: 0.013399352319538593
step: 360, loss: 0.01540173590183258
step: 370, loss: 0.1029503121972084
step: 380, loss: 0.01941845566034317
step: 390, loss: 0.1902616024017334
step: 400, loss: 0.2711779475212097
step: 410, loss: 0.04569073021411896
step: 420, loss: 0.05222201719880104
step: 430, loss: 0.17697520554065704
step: 440, loss: 0.020459400489926338
step: 450, loss: 0.02494249865412712
step: 460, loss: 0.007925021462142467
step: 470, loss: 0.15620636940002441
step: 480, loss: 0.10177326202392578
step: 490, loss: 0.08203937858343124
step: 500, loss: 0.0400925949215889
step: 510, loss: 0.09227853268384933
step: 520, loss: 0.025872599333524704
step: 530, loss: 0.01300827506929636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9284725426857406, f1=0.929784304726939,
best_f1=0.929784304726939
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

step: 0, loss: 0.2304660975933075
step: 10, loss: 0.0685761570930481
step: 20, loss: 0.01558742206543684
step: 30, loss: 0.11445004492998123
step: 40, loss: 0.26235324144363403
step: 50, loss: 0.14523017406463623
step: 60, loss: 0.012049800716340542
step: 70, loss: 0.015875011682510376
step: 80, loss: 0.060030169785022736
step: 90, loss: 0.01950003392994404
step: 100, loss: 0.03551284968852997
step: 110, loss: 0.11991959810256958
step: 120, loss: 0.26115697622299194
step: 130, loss: 0.11757072061300278
step: 140, loss: 0.032736774533987045
step: 150, loss: 0.10768896341323853
step: 160, loss: 0.04979144409298897
step: 170, loss: 0.010789585299789906
step: 180, loss: 0.031076889485120773
step: 190, loss: 0.042071886360645294
step: 200, loss: 0.00913113635033369
step: 210, loss: 0.07719338685274124
step: 220, loss: 0.05973225459456444
step: 230, loss: 0.01339353434741497
step: 240, loss: 0.008515733294188976
step: 250, loss: 0.050369344651699066
step: 260, loss: 0.0059548174031078815
step: 270, loss: 0.24851825833320618
step: 280, loss: 0.04491310939192772
step: 290, loss: 0.046347733587026596
step: 300, loss: 0.205389603972435
step: 310, loss: 0.02338734269142151
step: 320, loss: 0.10450409352779388
step: 330, loss: 0.03120938315987587
step: 340, loss: 0.16689518094062805
step: 350, loss: 0.012854074127972126
step: 360, loss: 0.10328326374292374
step: 370, loss: 0.11577589064836502
step: 380, loss: 0.0662461370229721
step: 390, loss: 0.08985821902751923
step: 400, loss: 0.023299764841794968
step: 410, loss: 0.011578240431845188
step: 420, loss: 0.22805839776992798
step: 430, loss: 0.01588333211839199
step: 440, loss: 0.19641900062561035
step: 450, loss: 0.025574522092938423
step: 460, loss: 0.03587519749999046
step: 470, loss: 0.03884318843483925

```

step: 480, loss: 0.22643645107746124
step: 490, loss: 0.02299177274107933
step: 500, loss: 0.21494153141975403
step: 510, loss: 0.04664485156536102
step: 520, loss: 0.05500565096735954
step: 530, loss: 0.03050341084599495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.931985294117647, f1=0.9251824817518249,
best_f1=0.9251824817518249
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03127370774745941
step: 10, loss: 0.06586199253797531
step: 20, loss: 0.0591057650744915
step: 30, loss: 0.022323554381728172
step: 40, loss: 0.010250689461827278
step: 50, loss: 0.11781751364469528
step: 60, loss: 0.03153245896100998
step: 70, loss: 0.01131503190845251
step: 80, loss: 0.0023442304227501154
step: 90, loss: 0.0032426936086267233
step: 100, loss: 0.023057634010910988
step: 110, loss: 0.11101873964071274
step: 120, loss: 0.0027141496539115906
step: 130, loss: 0.018843434751033783
step: 140, loss: 0.03356128931045532
step: 150, loss: 0.045808009803295135
step: 160, loss: 0.006419194396585226
step: 170, loss: 0.16395801305770874
step: 180, loss: 0.07537841796875
step: 190, loss: 0.08285857737064362
step: 200, loss: 0.06487192958593369
step: 210, loss: 0.08746829628944397
step: 220, loss: 0.060369718819856644
step: 230, loss: 0.03586912900209427
step: 240, loss: 0.0006958636804483831
step: 250, loss: 0.07518967986106873
step: 260, loss: 0.028020240366458893
step: 270, loss: 0.02511218748986721
step: 280, loss: 0.05930168554186821
step: 290, loss: 0.003947815857827663

```



```

step: 300, loss: 0.08592071384191513
step: 310, loss: 0.02940445765852928
step: 320, loss: 0.01096279639750719
step: 330, loss: 0.0011476161889731884
step: 340, loss: 0.013209376484155655
step: 350, loss: 0.004478665068745613
step: 360, loss: 0.014042233116924763
step: 370, loss: 0.017180873081088066
step: 380, loss: 0.010549076832830906
step: 390, loss: 0.004014675971120596
step: 400, loss: 0.02376568503677845
step: 410, loss: 0.004887694958597422
step: 420, loss: 0.12086282670497894
step: 430, loss: 0.013867199420928955
step: 440, loss: 0.023243283852934837
step: 450, loss: 0.0998539924621582
step: 460, loss: 0.033529628068208694
step: 470, loss: 0.010360165499150753
step: 480, loss: 0.0010671771597117186
step: 490, loss: 0.0009369294857606292
step: 500, loss: 0.014125674031674862
step: 510, loss: 0.0036030272021889687
step: 520, loss: 0.12428172677755356
step: 530, loss: 0.09726033359766006
/usr/local/lib/python3.7/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=0.9319029850746268, f1=0.9215686274509803,
best_f1=0.9251824817518249
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0151384761556983
step: 10, loss: 0.007702060975134373
step: 20, loss: 0.005878378637135029
step: 30, loss: 0.09944991767406464
step: 40, loss: 0.0015625189989805222
step: 50, loss: 0.07439842820167542
step: 60, loss: 0.07015606015920639
step: 70, loss: 0.0064657642506062984
step: 80, loss: 0.03256262466311455
step: 90, loss: 0.03556828573346138
step: 100, loss: 0.002672976115718484
step: 110, loss: 0.038312073796987534
step: 120, loss: 0.008631890639662743
step: 130, loss: 0.07048112899065018
step: 140, loss: 0.022393863648176193
step: 150, loss: 0.011579958721995354

```

```

step: 160, loss: 0.171139657497406
step: 170, loss: 0.008119868114590645
step: 180, loss: 0.004284133668988943
step: 190, loss: 0.021005189046263695
step: 200, loss: 0.003066832199692726
step: 210, loss: 0.12918400764465332
step: 220, loss: 0.003497564233839512
step: 230, loss: 0.0312797911465168
step: 240, loss: 0.008324399590492249
step: 250, loss: 0.013371913693845272
step: 260, loss: 0.06334331631660461
step: 270, loss: 0.015165876597166061
step: 280, loss: 0.0263985525816679
step: 290, loss: 0.03423236683011055
step: 300, loss: 0.0006211677100509405
step: 310, loss: 0.009418885223567486
step: 320, loss: 0.016956400126218796
step: 330, loss: 0.024582847952842712
step: 340, loss: 0.0011382800294086337
step: 350, loss: 0.002214202657341957
step: 360, loss: 0.017338665202260017
step: 370, loss: 0.003677277360111475
step: 380, loss: 0.010829341597855091
step: 390, loss: 0.00303968065418303
step: 400, loss: 0.04593709111213684
step: 410, loss: 0.004616255406290293
step: 420, loss: 0.00920820888131857
step: 430, loss: 0.05265449732542038
step: 440, loss: 0.0009293221519328654
step: 450, loss: 0.01039355993270874
step: 460, loss: 0.023498520255088806
step: 470, loss: 0.048531673848629
step: 480, loss: 0.009802197106182575
step: 490, loss: 0.006529544945806265
step: 500, loss: 0.014859583228826523
step: 510, loss: 0.10338962823152542
step: 520, loss: 0.09349282085895538
step: 530, loss: 0.0568724162876606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.935831381733021, f1=0.9289719626168225,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.036127980798482895
step: 10, loss: 0.041045181453228
step: 20, loss: 0.017612973228096962
step: 30, loss: 0.0012366812443360686
step: 40, loss: 0.07302258908748627
step: 50, loss: 0.002075514057651162
step: 60, loss: 0.01941888965666294
step: 70, loss: 0.0046392339281737804
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 80, loss: 0.000600749219302088
step: 90, loss: 0.12851563096046448
step: 100, loss: 0.00496671162545681
step: 110, loss: 0.0008080188999883831
step: 120, loss: 0.02621176280081272
step: 130, loss: 0.0012267122510820627
step: 140, loss: 0.009052220731973648
step: 150, loss: 0.030998162925243378
step: 160, loss: 0.0009712736937217414
step: 170, loss: 0.15267926454544067
step: 180, loss: 0.006268912926316261
step: 190, loss: 0.028393536806106567
step: 200, loss: 0.0023196435067802668
step: 210, loss: 0.012746225111186504
step: 220, loss: 0.0008689396199770272
step: 230, loss: 0.04277164116501808
step: 240, loss: 0.002788479905575514
step: 250, loss: 0.0027782537508755922
step: 260, loss: 0.013525744900107384
step: 270, loss: 0.000798641296569258
step: 280, loss: 0.019712507724761963
step: 290, loss: 0.1439356803894043
step: 300, loss: 0.006884533446282148
step: 310, loss: 0.01662001572549343
step: 320, loss: 0.10645108669996262
step: 330, loss: 0.0905006155371666
step: 340, loss: 0.004987109452486038
step: 350, loss: 0.0009866395266726613
step: 360, loss: 0.009240495041012764
step: 370, loss: 0.02035694010555744
step: 380, loss: 0.008872589096426964
step: 390, loss: 0.00032016393379308283
step: 400, loss: 0.002475622110068798
step: 410, loss: 0.03702891618013382
step: 420, loss: 0.00997175369411707
step: 430, loss: 0.015687275677919388
step: 440, loss: 0.006769683212041855
```

```

step: 450, loss: 0.0316438190639019
step: 460, loss: 0.00872889719903469
step: 470, loss: 0.0238113384693861
step: 480, loss: 0.015395567752420902
step: 490, loss: 0.0004327462229412049
step: 500, loss: 0.00835703406482935
step: 510, loss: 0.004693460650742054
step: 520, loss: 0.039681147783994675
step: 530, loss: 0.01890701614320278
/usr/local/lib/python3.7/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=0.9340761374187557, f1=0.9312906220984216,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002111943205818534
step: 10, loss: 0.0032453483436256647
step: 20, loss: 0.06285421550273895
step: 30, loss: 0.0031635835766792297
step: 40, loss: 0.0030857441015541553
step: 50, loss: 0.02152368612587452
step: 60, loss: 0.00016834333655424416
step: 70, loss: 0.0007830673130229115
step: 80, loss: 0.028236357495188713
step: 90, loss: 0.013763239607214928
step: 100, loss: 0.023113617673516273
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 110, loss: 0.001759466715157032
step: 120, loss: 0.28391000628471375
step: 130, loss: 0.00803354661911726
step: 140, loss: 0.013879876583814621
step: 150, loss: 0.035170093178749084
step: 160, loss: 0.014795220457017422
step: 170, loss: 0.0011932094348594546
step: 180, loss: 0.0008611801313236356
step: 190, loss: 0.0018343385308980942
step: 200, loss: 0.005719199310988188
step: 210, loss: 0.011809086427092552
step: 220, loss: 0.001907320343889296
step: 230, loss: 0.008516494184732437
step: 240, loss: 0.13321948051452637
step: 250, loss: 0.008891466073691845
step: 260, loss: 0.0032072202302515507
step: 270, loss: 0.09594354033470154
step: 280, loss: 0.0017021503299474716
step: 290, loss: 0.00021630791889037937

```

```

step: 300, loss: 0.06994830071926117
step: 310, loss: 0.005402350332587957
step: 320, loss: 0.001685724244453013
step: 330, loss: 0.0021728326100856066
step: 340, loss: 0.02288939617574215
step: 350, loss: 0.001553216832689941
step: 360, loss: 0.022073956206440926
step: 370, loss: 0.004527789540588856
step: 380, loss: 0.031337980180978775
step: 390, loss: 0.01721198670566082
step: 400, loss: 0.040384821593761444
step: 410, loss: 0.0010553275933489203
step: 420, loss: 0.036183856427669525
step: 430, loss: 0.0009403012227267027
step: 440, loss: 0.00030457269167527556
step: 450, loss: 0.0002022529224632308
step: 460, loss: 0.00047686105244793
step: 470, loss: 0.006437177304178476
step: 480, loss: 0.02414359711110592
step: 490, loss: 0.025370554998517036
step: 500, loss: 0.016739496961236
step: 510, loss: 0.015799010172486305
step: 520, loss: 0.001219455019959807
step: 530, loss: 0.10229156911373138
/usr/local/lib/python3.7/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=0.9263157894736842, f1=0.9223439733206289,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004820366739295423
step: 10, loss: 0.0016399382147938013
step: 20, loss: 0.007792723830789328
step: 30, loss: 0.009257500059902668
step: 40, loss: 0.06218644976615906
step: 50, loss: 0.026070410385727882
step: 60, loss: 0.03670414537191391
step: 70, loss: 0.02879066951572895
step: 80, loss: 0.02874823473393917
step: 90, loss: 0.002738968003541231
step: 100, loss: 0.00030260684434324503
step: 110, loss: 0.002068639500066638
step: 120, loss: 0.00661817193031311
step: 130, loss: 0.017142046242952347
step: 140, loss: 0.00010271491919411346
step: 150, loss: 0.10353396087884903

```

```

step: 160, loss: 0.0010418395977467299
step: 170, loss: 0.007615128066390753
step: 180, loss: 0.0012831324711441994
step: 190, loss: 0.0014570921193808317
step: 200, loss: 0.0011252970434725285
step: 210, loss: 0.0012988217640668154
step: 220, loss: 0.0003597074537537992
step: 230, loss: 0.008504454977810383
step: 240, loss: 0.000674059905577451
step: 250, loss: 0.0008131968206726015
step: 260, loss: 0.002367795677855611
step: 270, loss: 0.0022442527115345
step: 280, loss: 0.004333206452429295
step: 290, loss: 0.0028088660910725594
step: 300, loss: 0.005901915021240711
step: 310, loss: 0.000751662184484303
step: 320, loss: 0.0014129654737189412
step: 330, loss: 0.03173191472887993
step: 340, loss: 0.04084257408976555
step: 350, loss: 0.010305325500667095
step: 360, loss: 0.0067139589227736
step: 370, loss: 0.0048877340741455555
step: 380, loss: 0.0006414662930183113
step: 390, loss: 0.013727892190217972
step: 400, loss: 0.0004764254263136536
step: 410, loss: 0.0019989819265902042
step: 420, loss: 0.011847683228552341
step: 430, loss: 0.0018291858723387122
step: 440, loss: 0.07546535134315491
step: 450, loss: 0.0004071047296747565
step: 460, loss: 0.00032459417707286775
step: 470, loss: 0.004285715986043215
step: 480, loss: 0.10548784583806992
step: 490, loss: 0.003764424240216613
step: 500, loss: 0.013078613206744194
step: 510, loss: 0.13456493616104126
step: 520, loss: 0.004147139377892017
step: 530, loss: 0.004466884769499302
/usr/local/lib/python3.7/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=0.9308755760368663, f1=0.9337042188224385,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013968441635370255
step: 10, loss: 0.0012740673264488578

```

step: 20, loss: 0.0005568547639995813
step: 30, loss: 0.2843917906284332
step: 40, loss: 0.0001212691277032718
step: 50, loss: 0.009244505316019058
step: 60, loss: 0.007254383526742458
step: 70, loss: 0.0006638478953391314
step: 80, loss: 0.002205916214734316
step: 90, loss: 0.09376157820224762
step: 100, loss: 0.0014609259087592363
step: 110, loss: 0.0048766653053462505
step: 120, loss: 0.0006663592648692429
step: 130, loss: 0.0063848369754850864
step: 140, loss: 0.0008597875130362809
step: 150, loss: 0.0013957148184999824
step: 160, loss: 0.04443887621164322
step: 170, loss: 0.009083995595574379
step: 180, loss: 0.00011369732237653807
step: 190, loss: 0.0011295417789369822
step: 200, loss: 0.002174106426537037
step: 210, loss: 0.0015002256259322166
step: 220, loss: 0.0007287145708687603
step: 230, loss: 0.001157433376647532
step: 240, loss: 0.02515220455825329
step: 250, loss: 0.0005110138445161283
step: 260, loss: 0.0004191832267679274
step: 270, loss: 0.0011972261127084494
step: 280, loss: 0.28585585951805115
step: 290, loss: 0.0008106542518362403
step: 300, loss: 0.009808046743273735
step: 310, loss: 0.08076801151037216
step: 320, loss: 0.016513720154762268
step: 330, loss: 0.001455010031349957
step: 340, loss: 0.00010849506361410022
step: 350, loss: 0.0001848995016189292
step: 360, loss: 0.0029007818084210157
step: 370, loss: 0.0009999481262639165
step: 380, loss: 0.06760293990373611
step: 390, loss: 0.16116182506084442
step: 400, loss: 0.0013475967571139336
step: 410, loss: 0.04638671875
step: 420, loss: 0.005316935479640961
step: 430, loss: 0.05575443059206009
step: 440, loss: 0.031225116923451424
step: 450, loss: 0.000613710202742368
step: 460, loss: 0.0011444339761510491
step: 470, loss: 0.0008887294679880142
step: 480, loss: 0.00024813925847411156
step: 490, loss: 0.031403373926877975

```

step: 500, loss: 0.001649389392696321
step: 510, loss: 0.005562176462262869
step: 520, loss: 0.010710427537560463
step: 530, loss: 0.0005508763715624809
/usr/local/lib/python3.7/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=0.9350046425255338, f1=0.9330855018587362,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006608982221223414
step: 10, loss: 9.753812628332525e-05
step: 20, loss: 0.012566258199512959
step: 30, loss: 0.00031160638900473714
step: 40, loss: 0.00013254015357233584
step: 50, loss: 0.003708675503730774
step: 60, loss: 0.0001488519337726757
step: 70, loss: 0.003983202390372753
step: 80, loss: 0.0025048176757991314
step: 90, loss: 3.211131479474716e-05
step: 100, loss: 0.0033437085803598166
step: 110, loss: 0.00011017084761988372
step: 120, loss: 0.00030919822165742517
step: 130, loss: 0.043866414576768875
step: 140, loss: 0.0010142731480300426
step: 150, loss: 0.005809554364532232
step: 160, loss: 0.00033392017940059304
step: 170, loss: 0.013425000011920929
step: 180, loss: 0.00042993968236260116
step: 190, loss: 0.001063808100298047
step: 200, loss: 0.0003611628490034491
step: 210, loss: 0.00029011606238782406
step: 220, loss: 0.00026979329413734376
step: 230, loss: 0.0017583256121724844
step: 240, loss: 0.0005154876853339374
step: 250, loss: 0.0017658810829743743
step: 260, loss: 0.00991087406873703
step: 270, loss: 5.5047534260666e-05
step: 280, loss: 0.00041709517245180905
step: 290, loss: 0.027341941371560097
step: 300, loss: 0.00028129780548624694
step: 310, loss: 0.0008242553449235857
step: 320, loss: 0.0003146980598103255
step: 330, loss: 0.0006854363018646836
step: 340, loss: 0.0019205005373805761
step: 350, loss: 0.008824613876640797

```



```

step: 360, loss: 7.910741987871006e-05
step: 370, loss: 0.005943265277892351
step: 380, loss: 0.00030514583340846
step: 390, loss: 0.0001663332513999194
step: 400, loss: 0.002585864858701825
step: 410, loss: 0.00021202386415097862
step: 420, loss: 0.0006569541292265058
step: 430, loss: 0.0004294740210752934
step: 440, loss: 0.0026977595407515764
step: 450, loss: 0.0006005557952448726
step: 460, loss: 0.00012950006930623204
step: 470, loss: 0.00013715428940486163
step: 480, loss: 0.00016636328655295074
step: 490, loss: 0.0010840811301022768
step: 500, loss: 0.003165985457599163
step: 510, loss: 0.0038095058407634497
step: 520, loss: 0.0010932504665106535
step: 530, loss: 0.00011255505523877218
/usr/local/lib/python3.7/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=0.9250936329588015, f1=0.9225715626466449,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003396144602447748
step: 10, loss: 0.00011992715735686943
step: 20, loss: 0.00017847606795839965
step: 30, loss: 0.0011532288044691086
step: 40, loss: 0.00023766969388816506
step: 50, loss: 5.7765519159147516e-05
step: 60, loss: 0.0004696470859926194
step: 70, loss: 5.878507363377139e-05
step: 80, loss: 0.0001011321583064273
step: 90, loss: 0.00043970445403829217
step: 100, loss: 0.001943996292538941
step: 110, loss: 0.00014093273784965277
step: 120, loss: 5.2691364544443786e-05
step: 130, loss: 0.0005068217869848013
step: 140, loss: 0.0018053660169243813
step: 150, loss: 0.0003145506198052317
step: 160, loss: 0.0007515937322750688
step: 170, loss: 0.00010977408237522468
step: 180, loss: 8.23492300696671e-05
step: 190, loss: 0.0011733429273590446
step: 200, loss: 0.0004922430962324142
step: 210, loss: 0.0011706511722877622

```

```

step: 220, loss: 0.001898410846479237
step: 230, loss: 0.000531876168679446
step: 240, loss: 0.00021527787612285465
step: 250, loss: 0.00023305483045987785
step: 260, loss: 0.004223083145916462
step: 270, loss: 0.0015812565106898546
step: 280, loss: 0.028550706803798676
step: 290, loss: 0.00011967978934990242
step: 300, loss: 0.00010587961878627539
step: 310, loss: 0.000290554336970672
step: 320, loss: 0.0007926473044790328
step: 330, loss: 0.008853787556290627
step: 340, loss: 0.0002529959019739181
step: 350, loss: 0.05144836753606796
step: 360, loss: 0.0003444428148213774
step: 370, loss: 0.004371405579149723
step: 380, loss: 0.002275161212310195
step: 390, loss: 0.00013050738198217005
step: 400, loss: 0.0015995497815310955
step: 410, loss: 0.0017273633275181055
step: 420, loss: 0.001176983118057251
step: 430, loss: 0.00024424551520496607
step: 440, loss: 0.0007434709114022553
step: 450, loss: 0.0004207190650049597
step: 460, loss: 0.0002295111771672964
step: 470, loss: 0.00018963150796480477
step: 480, loss: 0.041733141988515854
step: 490, loss: 0.000375897012418136
step: 500, loss: 0.0496344156563282
step: 510, loss: 0.002978340955451131
step: 520, loss: 0.00138323032297194
step: 530, loss: 0.0008227488724514842
/usr/local/lib/python3.7/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=0.9257683215130024, f1=0.9161904761904762,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006183511111885309
step: 10, loss: 0.009700002148747444
step: 20, loss: 0.0004322330350987613
step: 30, loss: 0.0003445993934292346
step: 40, loss: 0.0011188415810465813
step: 50, loss: 0.022155988961458206
step: 60, loss: 0.0860411748290062
step: 70, loss: 0.00015604142390657216

```

step: 80, loss: 0.000468239450128749
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.000697278999723494
step: 100, loss: 0.03396039083600044
step: 110, loss: 0.015075547620654106
step: 120, loss: 0.0032000967767089605
step: 130, loss: 0.001421774853952229
step: 140, loss: 0.0007689374033361673
step: 150, loss: 0.00011166747572133318
step: 160, loss: 4.380721293273382e-05
step: 170, loss: 0.00012784163118340075
step: 180, loss: 0.00016433930431958288
step: 190, loss: 7.215381629066542e-05
step: 200, loss: 0.020556455478072166
step: 210, loss: 0.00043162042857147753
step: 220, loss: 0.02331605926156044
step: 230, loss: 0.0001416872109984979
step: 240, loss: 0.0008677492733113468
step: 250, loss: 0.001906550140120089
step: 260, loss: 0.0001671548088779673
step: 270, loss: 0.0007331918459385633
step: 280, loss: 0.00023865429102443159
step: 290, loss: 0.0005053990753367543
step: 300, loss: 7.844743231544271e-05
step: 310, loss: 0.00018134366837330163
step: 320, loss: 0.001343687530606985
step: 330, loss: 0.00031486907391808927
step: 340, loss: 5.546441752812825e-05
step: 350, loss: 5.879807940800674e-05
step: 360, loss: 5.198052531341091e-05
step: 370, loss: 0.0011367269326001406
step: 380, loss: 5.51946104678791e-05
step: 390, loss: 5.511147537617944e-05
step: 400, loss: 0.0009758605156093836
step: 410, loss: 0.0038225825410336256
step: 420, loss: 0.00014743870997335762
step: 430, loss: 8.37927291286178e-05
step: 440, loss: 0.000311432930175215
step: 450, loss: 0.00018823324353434145
step: 460, loss: 0.00018719356739893556
step: 470, loss: 3.715714774443768e-05
step: 480, loss: 0.0010311445221304893
step: 490, loss: 9.226264955941588e-05
step: 500, loss: 6.753928028047085e-05
step: 510, loss: 0.0002789102145470679
step: 520, loss: 0.0008386140689253807
step: 530, loss: 0.0026630021166056395
/usr/local/lib/python3.7/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=0.9292076887013595, f1=0.9235100891600189,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004140136763453484
step: 10, loss: 3.693836333695799e-05
step: 20, loss: 0.012024296447634697
step: 30, loss: 0.0001530235167592764
step: 40, loss: 0.0008427543216384947
step: 50, loss: 0.0064666904509067535
step: 60, loss: 0.002674748655408621
step: 70, loss: 0.00011237413855269551
step: 80, loss: 0.00016834016423672438
step: 90, loss: 0.001277645817026496
step: 100, loss: 0.00022549911227542907
step: 110, loss: 8.112001523841172e-05
step: 120, loss: 8.962200081441551e-05
step: 130, loss: 0.0017346425447613
step: 140, loss: 0.02009095624089241
step: 150, loss: 0.0002722779172472656
step: 160, loss: 0.00028993390151299536
step: 170, loss: 0.0003728807205334306
step: 180, loss: 0.00010088331328006461
step: 190, loss: 0.0005703077185899019
step: 200, loss: 0.0005429785815067589
step: 210, loss: 0.00044286984484642744
step: 220, loss: 6.166548701003194e-05
step: 230, loss: 6.280246452661231e-05
step: 240, loss: 0.0009836372919380665
step: 250, loss: 0.0017469319282099605
step: 260, loss: 0.0001132977704401128
step: 270, loss: 9.156251326203346e-05
step: 280, loss: 0.027523312717676163
step: 290, loss: 0.001828458858653903
step: 300, loss: 0.005202615633606911
step: 310, loss: 0.00036504241870716214
step: 320, loss: 5.8888603234663606e-05
step: 330, loss: 0.0001548353029647842
step: 340, loss: 0.00032244520843960345
step: 350, loss: 0.0002857298240996897
step: 360, loss: 0.00016083900118246675
step: 370, loss: 0.0004252833023201674
step: 380, loss: 0.0004267265321686864
step: 390, loss: 0.0022220283281058073
step: 400, loss: 0.00029382313368842006
```

```

step: 410, loss: 0.005697300657629967
step: 420, loss: 0.03493368625640869
step: 430, loss: 0.002300616819411516
step: 440, loss: 0.004414400551468134
step: 450, loss: 0.0007699363050051033
step: 460, loss: 9.502691682428122e-05
step: 470, loss: 9.541920735500753e-05
step: 480, loss: 0.0021321100648492575
step: 490, loss: 0.00035032874438911676
step: 500, loss: 0.00015108443039935082
step: 510, loss: 0.0018416864331811666
step: 520, loss: 3.665813710540533e-05
step: 530, loss: 0.055099062621593475
/usr/local/lib/python3.7/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=0.9284712482468443, f1=0.9256820319849483,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002093459013849497
step: 10, loss: 0.0002484804717823863
step: 20, loss: 0.003931957762688398
step: 30, loss: 0.004539126064628363
step: 40, loss: 0.00015848023758735508
step: 50, loss: 0.0004795166023541242
step: 60, loss: 0.0002486702287569642
step: 70, loss: 8.403797255596146e-05
step: 80, loss: 0.00023611244978383183
step: 90, loss: 0.012485766783356667
step: 100, loss: 0.0006204357487149537
step: 110, loss: 0.002663047518581152
step: 120, loss: 0.005804837681353092
step: 130, loss: 0.0009138752357102931
step: 140, loss: 0.03051638789474964
step: 150, loss: 0.0016841395990923047
step: 160, loss: 4.235333108226769e-05
step: 170, loss: 0.00022314822126645595
step: 180, loss: 8.979316771728918e-05
step: 190, loss: 0.0002728084218688309
step: 200, loss: 9.350771142635494e-05
step: 210, loss: 0.0007927849073894322
step: 220, loss: 9.051791857928038e-05
step: 230, loss: 0.005404164083302021
step: 240, loss: 0.0007669439655728638
step: 250, loss: 0.0004247694159857929
step: 260, loss: 0.0012273866450414062

```

```

step: 270, loss: 0.00011459668894531205
step: 280, loss: 0.00027331410092301667
step: 290, loss: 0.001853934838436544
step: 300, loss: 6.405235035344958e-05
step: 310, loss: 0.0037577508483082056
step: 320, loss: 0.006621803622692823
step: 330, loss: 0.0007742154411971569
step: 340, loss: 0.0003310991742182523
step: 350, loss: 0.0006882966263219714
step: 360, loss: 5.040665564592928e-05
step: 370, loss: 0.002101023681461811
step: 380, loss: 0.000152743115904741
step: 390, loss: 0.0062352982349693775
step: 400, loss: 0.000329487316776067
step: 410, loss: 0.0014702138723805547
step: 420, loss: 0.00024486170150339603
step: 430, loss: 0.0003811752831097692
step: 440, loss: 0.00017332231800537556
step: 450, loss: 9.668127313489094e-05
step: 460, loss: 0.0005924534052610397
step: 470, loss: 0.00014015917258802801
step: 480, loss: 0.00010788116924231872
step: 490, loss: 0.003309716237708926
step: 500, loss: 0.00037394592072814703
step: 510, loss: 0.004057477228343487
step: 520, loss: 0.0003109322569798678
step: 530, loss: 0.0002447979059070349
/usr/local/lib/python3.7/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=0.9301895515487749, f1=0.9264229523368811,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008828625432215631
step: 10, loss: 3.681459202198312e-05
step: 20, loss: 0.00013902199862059206
step: 30, loss: 6.671086885035038e-05
step: 40, loss: 0.00019689188047777861
step: 50, loss: 0.00862960983067751
step: 60, loss: 0.00044714572140946984
step: 70, loss: 0.0002905844885390252
step: 80, loss: 2.9506347345886752e-05
step: 90, loss: 0.0007768190116621554
step: 100, loss: 0.00017259051674045622
step: 110, loss: 9.704477997729555e-05
step: 120, loss: 0.00033363429247401655

```

```

step: 130, loss: 0.006469429470598698
step: 140, loss: 0.0003493014955893159
step: 150, loss: 0.00047237929538823664
step: 160, loss: 0.0004237140528857708
step: 170, loss: 0.00041165947914123535
step: 180, loss: 5.441347093437798e-05
step: 190, loss: 0.00036625683424063027
step: 200, loss: 0.0001330904196947813
step: 210, loss: 4.0920036553870887e-05
step: 220, loss: 0.0025351541116833687
step: 230, loss: 0.000347140128724277
step: 240, loss: 0.00017475885397288948
step: 250, loss: 0.0034296808298677206
step: 260, loss: 0.0016037391033023596
step: 270, loss: 0.00017077724623959512
step: 280, loss: 0.00019122511730529368
step: 290, loss: 0.007198140025138855
step: 300, loss: 0.002270712750032544
step: 310, loss: 0.000772113271523267
step: 320, loss: 0.0005500062834471464
step: 330, loss: 0.0020467224530875683
step: 340, loss: 0.00019904789223801345
step: 350, loss: 0.00033528023050166667
step: 360, loss: 0.005053210072219372
step: 370, loss: 9.535243589198217e-05
step: 380, loss: 0.0006410346832126379
step: 390, loss: 0.035266581922769547
step: 400, loss: 0.00027669465634971857
step: 410, loss: 0.002117678988724947
step: 420, loss: 0.23341630399227142
step: 430, loss: 4.383133273222484e-05
step: 440, loss: 0.006542154587805271
step: 450, loss: 0.007428754586726427
step: 460, loss: 0.0003629977290984243
step: 470, loss: 0.00032317920704372227
step: 480, loss: 0.0001225252781296149
step: 490, loss: 3.832978109130636e-05
step: 500, loss: 0.0003525686333887279
step: 510, loss: 0.008421283215284348
step: 520, loss: 0.0004905915702693164
step: 530, loss: 0.00812464114278555
/usr/local/lib/python3.7/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=0.9329556185080263, f1=0.9248937175247992,
best_f1=0.9289719626168225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.638316411525011e-05
step: 10, loss: 0.00013386573118623346
step: 20, loss: 0.000956860720179975
step: 30, loss: 0.00015131205145735294
step: 40, loss: 0.02402821183204651
step: 50, loss: 8.263030758826062e-05
step: 60, loss: 0.000137067268951796
step: 70, loss: 0.0014191848458722234
step: 80, loss: 7.646105950698256e-05
step: 90, loss: 0.00010268065670970827
step: 100, loss: 0.0003870357759296894
step: 110, loss: 0.0003044807817786932
step: 120, loss: 0.0005781413055956364
step: 130, loss: 0.132158562541008
step: 140, loss: 0.00040255545172840357
step: 150, loss: 6.30418217042461e-05
step: 160, loss: 0.0001228264154633507
step: 170, loss: 0.0001469750568503514
step: 180, loss: 6.435428804252297e-05
step: 190, loss: 0.00012690799485426396
step: 200, loss: 0.00013557725469581783
step: 210, loss: 0.01124142948538065
step: 220, loss: 3.606607424444519e-05
step: 230, loss: 0.10171453654766083
step: 240, loss: 5.560824502026662e-05
step: 250, loss: 0.00024255149764940143
step: 260, loss: 8.730378613108769e-05
step: 270, loss: 2.540941750339698e-05
step: 280, loss: 5.254785719444044e-05
step: 290, loss: 2.353581294300966e-05
step: 300, loss: 2.106990541506093e-05
step: 310, loss: 0.0003605015517678112
step: 320, loss: 7.199015817604959e-05
step: 330, loss: 0.0012306615244597197
step: 340, loss: 7.128303695935756e-05
step: 350, loss: 8.047474693739787e-05
step: 360, loss: 0.0005629633087664843
step: 370, loss: 0.000157530652359128
step: 380, loss: 9.745806164573878e-05
step: 390, loss: 0.0031704904977232218
step: 400, loss: 0.01190404873341322
step: 410, loss: 0.0005819576908834279
step: 420, loss: 0.0011317136231809855
step: 430, loss: 0.00014207027561496943
step: 440, loss: 0.00042443061829544604
step: 450, loss: 4.756827183882706e-05
step: 460, loss: 2.254839455417823e-05
```



```

step: 470, loss: 0.0003973132697865367
step: 480, loss: 3.802841456490569e-05
step: 490, loss: 2.1877840481465682e-05
step: 500, loss: 0.030920125544071198
step: 510, loss: 0.00010846649820450693
step: 520, loss: 3.428575291763991e-05
step: 530, loss: 0.00018212015856988728
/usr/local/lib/python3.7/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=0.9328891004283675, f1=0.9235154394299286,
best_f1=0.9289719626168225
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 [00:29, 196.70it/s]
load_f1 = 0.9339578454332552
real_f1 = 0.9322671683913453
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:22, 192.63it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5628790259361267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5161290322580646, f1=0.4, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.46800440549850464
/usr/local/lib/python3.7/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=0.4615384615384615, f1=0.39285714285714285, best_f1=0.4
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4539681077003479

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6086956521739131, f1=0.41025641025641024,
best_f1=0.41025641025641024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3242473602294922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7096774193548386, f1=0.4782608695652174,
best_f1=0.4782608695652174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23497581481933594
/usr/local/lib/python3.7/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=0.5263157894736842, f1=0.64, best_f1=0.4782608695652174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3422366976737976
/usr/local/lib/python3.7/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=0.6, f1=0.5833333333333334, best_f1=0.4782608695652174
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19782574474811554
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8, f1=0.6250000000000001, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011447827331721783
/usr/local/lib/python3.7/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=0.8, f1=0.6250000000000001, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026455437764525414
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6250000000000001,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006094394717365503
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.6666666666666666,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027552435640245676
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6666666666666666,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00666341558098793
/usr/local/lib/python3.7/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=0.8, f1=0.6206896551724138, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.006637635175138712
/usr/local/lib/python3.7/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=0.782608695652174, f1=0.5714285714285714,
best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003047654405236244
/usr/local/lib/python3.7/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=0.8, f1=0.6206896551724138, best_f1=0.6250000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006738625466823578
/usr/local/lib/python3.7/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=0.8, f1=0.6206896551724138, best_f1=0.6250000000000001
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, 139605.58it/s]
/usr/local/lib/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 = 0.8
real_f1 = 0.7407407407407408
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, 190.03it/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 bert-base-uncased \
    --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 bert-base-uncased \
    --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: 0.6499109268188477
step: 10, loss: 0.6404380798339844
step: 20, loss: 0.33818355202674866
step: 30, loss: 0.10740054398775101
step: 40, loss: 0.43571341037750244
step: 50, loss: 0.0387084074318409
step: 60, loss: 0.0304944459348917
step: 70, loss: 0.0627317875623703
step: 80, loss: 0.07253143936395645
step: 90, loss: 0.24622781574726105
step: 100, loss: 0.005992227699607611
step: 110, loss: 0.2653926610946655
step: 120, loss: 0.01825672574341297
step: 130, loss: 0.006490346975624561
step: 140, loss: 0.004942466039210558
step: 150, loss: 0.007207461632788181

```



```

step: 160, loss: 0.006130199879407883
step: 170, loss: 0.10741986334323883
step: 180, loss: 0.019619911909103394
step: 190, loss: 0.0998012125492096
step: 200, loss: 0.019376881420612335
step: 210, loss: 0.0052696396596729755
step: 220, loss: 0.002717965515330434
step: 230, loss: 0.2562961280345917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9820627802690582, f1=0.9750566893424036,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028902951162308455
step: 10, loss: 0.0030420708935707808
step: 20, loss: 0.08488193899393082
step: 30, loss: 0.17557646334171295
step: 40, loss: 0.05593681335449219
step: 50, loss: 0.042870309203863144
step: 60, loss: 0.0028662518598139286
step: 70, loss: 0.11992425471544266
step: 80, loss: 0.0074059502221643925
step: 90, loss: 0.14768856763839722
step: 100, loss: 0.16069245338439941
step: 110, loss: 0.027385877445340157
step: 120, loss: 0.13472767174243927
step: 130, loss: 0.025348249822854996
step: 140, loss: 0.005405582953244448
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.005307094659656286
step: 160, loss: 0.0020937644876539707
step: 170, loss: 0.001976720057427883
step: 180, loss: 0.001152555807493627
step: 190, loss: 0.007001284975558519
step: 200, loss: 0.005872141104191542
step: 210, loss: 0.0005663957563228905
step: 220, loss: 0.04687467962503433
step: 230, loss: 0.16096359491348267
/usr/local/lib/python3.7/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=0.9730941704035874, f1=0.9761092150170648,
best_f1=0.9750566893424036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026318546384572983
step: 10, loss: 0.013838320970535278
step: 20, loss: 0.0017308200476691127
step: 30, loss: 0.013464020565152168
step: 40, loss: 0.04790053516626358
step: 50, loss: 0.00860634446144104
step: 60, loss: 0.03948524594306946
step: 70, loss: 0.010546554811298847
step: 80, loss: 0.000841242668684572
step: 90, loss: 0.1957041621208191
step: 100, loss: 0.005007543601095676
step: 110, loss: 0.0007771142991259694
step: 120, loss: 0.051020704209804535
step: 130, loss: 0.006505885161459446
step: 140, loss: 0.008034494705498219
step: 150, loss: 0.003458187449723482
step: 160, loss: 0.004795626271516085
step: 170, loss: 0.010964491404592991
step: 180, loss: 0.0062615820206701756
step: 190, loss: 0.002107849810272455
step: 200, loss: 0.016472330316901207
step: 210, loss: 0.02824213169515133
step: 220, loss: 0.0005137564148753881
step: 230, loss: 0.001352141727693379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9764309764309763,
best_f1=0.9764309764309763
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006390249473042786
step: 10, loss: 0.0002836911880876869
step: 20, loss: 0.002010715426877141
step: 30, loss: 0.0005628179642371833
step: 40, loss: 0.003220637096092105
step: 50, loss: 0.0007250572089105844
step: 60, loss: 0.030013099312782288

```

```

step: 70, loss: 0.00834976602345705
step: 80, loss: 0.0010333952959626913
step: 90, loss: 0.0012954360572621226
step: 100, loss: 0.000581770611461252
step: 110, loss: 0.0007268697954714298
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.08803177624940872
step: 130, loss: 0.033187732100486755
step: 140, loss: 0.001997441751882434
step: 150, loss: 0.004102783277630806
step: 160, loss: 0.0002737884351518005
step: 170, loss: 0.0027118981815874577
step: 180, loss: 0.0009003353770822287
step: 190, loss: 0.0008849338046275079
step: 200, loss: 0.0024761618115007877
step: 210, loss: 0.0018983418121933937
step: 220, loss: 0.0002587438502814621
step: 230, loss: 0.005022658035159111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9887387387387387, f1=0.9819413092550789,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00034429033985361457
step: 10, loss: 0.0003005805774591863
step: 20, loss: 0.0009440059075132012
step: 30, loss: 0.00022006318613421172
step: 40, loss: 0.0004901050706394017
step: 50, loss: 0.0005131374346092343
step: 60, loss: 0.00132584641687572
step: 70, loss: 0.001550248940475285
step: 80, loss: 0.0007385800709016621
step: 90, loss: 0.000344013242283836
step: 100, loss: 0.00021505825861822814
step: 110, loss: 0.0005877810763195157
step: 120, loss: 8.567688928451389e-05
step: 130, loss: 0.00031277057132683694
step: 140, loss: 0.0036069871857762337
step: 150, loss: 0.0011335948947817087
step: 160, loss: 0.0006685375119559467
step: 170, loss: 0.051312293857336044

```

```

step: 180, loss: 0.00833399873226881
step: 190, loss: 0.0076180389150977135
step: 200, loss: 0.022601531818509102
step: 210, loss: 0.001622588373720646
step: 220, loss: 0.0029344195500016212
step: 230, loss: 0.0006067244685254991
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9831649831649831,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001258359756320715
step: 10, loss: 0.0005805341061204672
step: 20, loss: 0.0035291300155222416
step: 30, loss: 0.0007948352722451091
step: 40, loss: 0.00026177382096648216
step: 50, loss: 0.00023698896984569728
step: 60, loss: 0.00014981160347815603
step: 70, loss: 0.00038707340718246996
step: 80, loss: 0.0013755386462435126
step: 90, loss: 0.10804078727960587
step: 100, loss: 0.032018497586250305
step: 110, loss: 0.0025057268794625998
step: 120, loss: 0.0030430436599999666
step: 130, loss: 0.0011317646130919456
step: 140, loss: 0.00014395129983313382
step: 150, loss: 0.00031024092459119856
step: 160, loss: 0.005494014360010624
step: 170, loss: 0.00015672114386688918
step: 180, loss: 0.002560992958024144
step: 190, loss: 0.03642645478248596
step: 200, loss: 0.0002681954065337777
step: 210, loss: 0.0007197900558821857
step: 220, loss: 0.0003659986541606486
step: 230, loss: 0.044792819768190384
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9785794813979707,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015267760027199984
step: 10, loss: 0.0001697596162557602

```

```

step: 20, loss: 0.001133991521783173
step: 30, loss: 0.016385924071073532
step: 40, loss: 0.0006883738678880036
step: 50, loss: 0.0023778830654919147
step: 60, loss: 0.00089441635645926
step: 70, loss: 0.01443660631775856
step: 80, loss: 0.00043627448030747473
step: 90, loss: 0.00010141950770048425
step: 100, loss: 0.000171036968822591
step: 110, loss: 0.0004819768655579537
step: 120, loss: 0.00025616015773266554
step: 130, loss: 0.0006762737757526338
step: 140, loss: 0.0002930292102973908
step: 150, loss: 0.005617765709757805
step: 160, loss: 0.051642756909132004
step: 170, loss: 0.0006658969214186072
step: 180, loss: 0.0017408514395356178
step: 190, loss: 0.0003176049212925136
step: 200, loss: 0.031610533595085144
step: 210, loss: 0.00014886818826198578
step: 220, loss: 0.004541992675513029
step: 230, loss: 0.0035705030895769596
/usr/local/lib/python3.7/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=0.9863636363636363, f1=0.9806598407281,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013374965637922287
step: 10, loss: 0.00047512276796624064
step: 20, loss: 0.00011524009460117668
step: 30, loss: 0.00016173605399671942
step: 40, loss: 0.000538735359441489
step: 50, loss: 0.00023987094755284488
step: 60, loss: 0.0005943485884927213
step: 70, loss: 0.00023697350115980953
step: 80, loss: 0.0021772868931293488
step: 90, loss: 0.00018048447964247316
step: 100, loss: 0.0008017154177650809
step: 110, loss: 0.0013392710825428367
step: 120, loss: 0.00022335625544656068
step: 130, loss: 0.007440399844199419
step: 140, loss: 0.0001825473882490769
step: 150, loss: 0.00030576949939131737
step: 160, loss: 0.0009877369739115238
step: 170, loss: 0.00026091362815350294

```

```

step: 180, loss: 9.741173562360927e-05
step: 190, loss: 0.00018503218598198146
step: 200, loss: 0.00017073056369554251
step: 210, loss: 0.00012707685527857393
step: 220, loss: 0.0003257939824834466
step: 230, loss: 0.00017310735711362213
/usr/local/lib/python3.7/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=0.9875424688561721, f1=0.9772727272727272,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011777962208725512
step: 10, loss: 0.0006153172580525279
step: 20, loss: 0.0007745450711809099
step: 30, loss: 6.51234295219183e-05
step: 40, loss: 0.03178218752145767
step: 50, loss: 0.00012249067367520183
step: 60, loss: 0.00011851927411044016
step: 70, loss: 0.001919787609949708
step: 80, loss: 0.0010732188820838928
step: 90, loss: 0.00020442645472940058
step: 100, loss: 0.00026922710821963847
step: 110, loss: 0.001838106196373701
step: 120, loss: 9.186736861011013e-05
step: 130, loss: 4.312100281822495e-05
step: 140, loss: 4.839895336772315e-05
step: 150, loss: 0.00024794391356408596
step: 160, loss: 7.037647446850315e-05
step: 170, loss: 0.00010992435272783041
step: 180, loss: 0.0018144132336601615
step: 190, loss: 0.00010586456482997164
step: 200, loss: 6.798967660870403e-05
step: 210, loss: 0.0007703670999035239
step: 220, loss: 4.9479414883535355e-05
step: 230, loss: 0.0001025608362397179
/usr/local/lib/python3.7/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=0.9864253393665158, f1=0.9807037457434733,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.266698569059372e-05
step: 10, loss: 2.7864369258168153e-05

```

```

step: 20, loss: 7.076525798765942e-05
step: 30, loss: 0.0001288107450818643
step: 40, loss: 0.0017209912184625864
step: 50, loss: 0.0001353437837678939
step: 60, loss: 0.001457109465263784
step: 70, loss: 0.0026412836741656065
step: 80, loss: 9.239283826900646e-05
step: 90, loss: 0.0003521083854138851
step: 100, loss: 0.00017213723913300782
step: 110, loss: 8.267957309726626e-05
step: 120, loss: 8.474523929180577e-05
step: 130, loss: 7.986660784808919e-05
step: 140, loss: 0.06927884370088577
step: 150, loss: 0.009639851748943329
step: 160, loss: 5.55875594727695e-05
step: 170, loss: 0.0008785890531726182
step: 180, loss: 7.435668521793559e-05
step: 190, loss: 0.014198931865394115
step: 200, loss: 0.004320110194385052
step: 210, loss: 5.6249515182571486e-05
step: 220, loss: 0.0001812630071071908
step: 230, loss: 0.002798826899379492
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9865470852017937, f1=0.9831271091113611,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.470128523185849e-05
step: 10, loss: 0.00023333601711783558
step: 20, loss: 0.0008139958954416215
step: 30, loss: 0.025641456246376038
step: 40, loss: 0.00018291579908691347
step: 50, loss: 0.00011145273310830817
step: 60, loss: 9.51944530243054e-05
step: 70, loss: 0.0004295269900467247
step: 80, loss: 4.322562017478049e-05
step: 90, loss: 6.999829201959074e-05
step: 100, loss: 6.21319777565077e-05
step: 110, loss: 0.014245384372770786
step: 120, loss: 4.8328387492801994e-05
step: 130, loss: 3.3045304007828236e-05
step: 140, loss: 5.17582411703188e-05
step: 150, loss: 0.024614961817860603
step: 160, loss: 4.5561155275208876e-05
step: 170, loss: 0.023466214537620544

```

```

step: 180, loss: 6.803734868299216e-05
step: 190, loss: 6.0815702454419807e-05
step: 200, loss: 9.645529644330963e-05
step: 210, loss: 4.0472830733051524e-05
step: 220, loss: 6.148131069494411e-05
step: 230, loss: 0.00012699011131189764
/usr/local/lib/python3.7/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=0.9875706214689265, f1=0.9806598407281,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.726773088099435e-05
step: 10, loss: 4.803081537829712e-05
step: 20, loss: 4.558289219858125e-05
step: 30, loss: 7.004579674685374e-05
step: 40, loss: 9.123428026214242e-05
step: 50, loss: 0.00020978820975869894
step: 60, loss: 0.0002706437953747809
step: 70, loss: 6.016105544404127e-05
step: 80, loss: 8.107251778710634e-05
step: 90, loss: 5.808902278658934e-05
step: 100, loss: 0.0001347788784187287
step: 110, loss: 8.746454113861546e-05
step: 120, loss: 6.600791675737128e-05
step: 130, loss: 4.970364534528926e-05
step: 140, loss: 0.013006370514631271
step: 150, loss: 0.00010073917655972764
step: 160, loss: 6.310873868642375e-05
step: 170, loss: 6.0834918258478865e-05
step: 180, loss: 0.024975404143333435
step: 190, loss: 3.769757677218877e-05
step: 200, loss: 2.7897362087969668e-05
step: 210, loss: 5.3642950661014766e-05
step: 220, loss: 5.592870365944691e-05
step: 230, loss: 0.07343403995037079
/usr/local/lib/python3.7/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=0.9864253393665158, f1=0.979591836734694,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.765325426589698e-05
step: 10, loss: 0.00021181110059842467

```



```

step: 20, loss: 0.0004999861703254282
step: 30, loss: 7.976012420840561e-05
step: 40, loss: 0.0001394561113556847
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.0001325777411693707
step: 60, loss: 5.178566061658785e-05
step: 70, loss: 3.0419440008699894e-05
step: 80, loss: 4.4566153519554064e-05
step: 90, loss: 3.996642408310436e-05
step: 100, loss: 4.187223385088146e-05
step: 110, loss: 0.0262893158942461
step: 120, loss: 0.0443478487432003
step: 130, loss: 6.082091204007156e-05
step: 140, loss: 4.501776857068762e-05
step: 150, loss: 6.026994378771633e-05
step: 160, loss: 9.220552601618692e-05
step: 170, loss: 6.262618990149349e-05
step: 180, loss: 5.088022953714244e-05
step: 190, loss: 3.386514072190039e-05
step: 200, loss: 6.828460027463734e-05
step: 210, loss: 2.78157622233266e-05
step: 220, loss: 5.151160803507082e-05
step: 230, loss: 4.9019119614968076e-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=0.9875141884222476, f1=0.9794520547945206,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.025319867650978e-05
step: 10, loss: 0.00981968268752098
step: 20, loss: 5.0379465392325073e-05
step: 30, loss: 7.64135256758891e-05
step: 40, loss: 0.00016172986943274736
step: 50, loss: 4.4712269300362095e-05
step: 60, loss: 6.944490451132879e-05
step: 70, loss: 6.178942567203194e-05
step: 80, loss: 5.0298905989620835e-05
step: 90, loss: 0.00011610796354943886
step: 100, loss: 4.775624620378949e-05
step: 110, loss: 5.1423041441012174e-05
step: 120, loss: 2.6928828447125852e-05
step: 130, loss: 5.040309770265594e-05
step: 140, loss: 7.106881821528077e-05
step: 150, loss: 4.130293382331729e-05
step: 160, loss: 0.003184210741892457

```

```

step: 170, loss: 1.9140275981044397e-05
step: 180, loss: 3.430474316701293e-05
step: 190, loss: 0.00010257851681672037
step: 200, loss: 3.27958186971955e-05
step: 210, loss: 4.8828420403879136e-05
step: 220, loss: 9.502867760602385e-05
step: 230, loss: 9.120386675931513e-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=0.9875706214689265, f1=0.9830124575311437,
best_f1=0.9819413092550789
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.199849743396044e-05
step: 10, loss: 2.1829562683706172e-05
step: 20, loss: 4.5341712393565103e-05
step: 30, loss: 4.117372373002581e-05
step: 40, loss: 5.9097168559674174e-05
step: 50, loss: 0.05776557698845863
step: 60, loss: 4.2954226955771446e-05
step: 70, loss: 0.0021005484741181135
step: 80, loss: 0.0009846960892900825
step: 90, loss: 0.0002371605223743245
step: 100, loss: 5.030544343753718e-05
step: 110, loss: 4.235721644363366e-05
step: 120, loss: 6.512441905215383e-05
step: 130, loss: 0.018132973462343216
step: 140, loss: 3.27550878864713e-05
step: 150, loss: 0.00024897779803723097
step: 160, loss: 4.168817395111546e-05
step: 170, loss: 8.500654075760394e-05
step: 180, loss: 0.0005866176798008382
step: 190, loss: 0.0009046823251992464
step: 200, loss: 5.7410226872889325e-05
step: 210, loss: 3.333642234792933e-05
step: 220, loss: 3.152180215693079e-05
step: 230, loss: 3.5973102058051154e-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=0.9875706214689265, f1=0.9830124575311437,
best_f1=0.9819413092550789
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.")
2473it [00:15, 160.90it/s]
load_f1 = 0.9876265466816648
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 [00:24, 183.01it/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 bert-base-uncased \
--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 bert-base-uncased \
  --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: 0.6259139180183411
step: 10, loss: 0.5429301261901855
step: 20, loss: 0.5519333481788635
step: 30, loss: 0.2263692170381546
step: 40, loss: 0.2085190862417221
step: 50, loss: 0.2415197640657425
step: 60, loss: 0.09252595156431198
step: 70, loss: 0.10910136997699738
step: 80, loss: 0.12536096572875977
step: 90, loss: 0.4425000250339508
step: 100, loss: 0.045878905802965164
step: 110, loss: 0.07744130492210388
step: 120, loss: 0.11383363604545593
step: 130, loss: 0.15734638273715973
step: 140, loss: 0.15088671445846558
step: 150, loss: 0.057592831552028656
step: 160, loss: 0.03142508491873741
step: 170, loss: 0.25414931774139404
step: 180, loss: 0.020348068326711655
step: 190, loss: 0.02034829370677471
step: 200, loss: 0.1573685258626938
step: 210, loss: 0.08803804218769073
step: 220, loss: 0.36695489287376404
step: 230, loss: 0.1698952466249466
step: 240, loss: 0.06353004276752472
step: 250, loss: 0.021902363747358322
step: 260, loss: 0.10909101366996765
step: 270, loss: 0.005421613808721304
step: 280, loss: 0.029140204191207886
step: 290, loss: 0.24269813299179077
step: 300, loss: 0.08072756975889206
step: 310, loss: 0.35138827562332153
step: 320, loss: 0.056562453508377075
step: 330, loss: 0.1376778930425644
step: 340, loss: 0.07603613287210464
step: 350, loss: 0.04859049245715141
step: 360, loss: 0.03328530862927437
step: 370, loss: 0.1221434473991394
step: 380, loss: 0.015347711741924286
step: 390, loss: 0.20221586525440216
```

```

step: 400, loss: 0.3347606658935547
step: 410, loss: 0.03526876121759415
step: 420, loss: 0.03415589779615402
step: 430, loss: 0.17614834010601044
step: 440, loss: 0.013476529158651829
step: 450, loss: 0.009422343224287033
step: 460, loss: 0.009144806303083897
step: 470, loss: 0.08354010432958603
step: 480, loss: 0.041020315140485764
step: 490, loss: 0.07238859683275223
step: 500, loss: 0.2213204950094223
step: 510, loss: 0.04803695157170296
step: 520, loss: 0.07377597689628601
step: 530, loss: 0.003504390362650156
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.922166590805644, f1=0.9253731343283582,
best_f1=0.9253731343283582
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11322378367185593
step: 10, loss: 0.07562016695737839
step: 20, loss: 0.04994134604930878
step: 30, loss: 0.03877311572432518
step: 40, loss: 0.08610009402036667
step: 50, loss: 0.2055002599954605
step: 60, loss: 0.023130353540182114
step: 70, loss: 0.02336384542286396
step: 80, loss: 0.05858834832906723
step: 90, loss: 0.04307268559932709
step: 100, loss: 0.02351577766239643
step: 110, loss: 0.05777204409241676
step: 120, loss: 0.06714864075183868
step: 130, loss: 0.12024373561143875
step: 140, loss: 0.05106078088283539
step: 150, loss: 0.06278771162033081
step: 160, loss: 0.020358623936772346
step: 170, loss: 0.04242348298430443
step: 180, loss: 0.038977138698101044
step: 190, loss: 0.13501864671707153
step: 200, loss: 0.01954326033592224
step: 210, loss: 0.05230044201016426

```

```

step: 220, loss: 0.07767936587333679
step: 230, loss: 0.009143305942416191
step: 240, loss: 0.0353456474840641
step: 250, loss: 0.09594839811325073
step: 260, loss: 0.0032147967722266912
step: 270, loss: 0.24113714694976807
step: 280, loss: 0.03043900430202484
step: 290, loss: 0.0479779876768589
step: 300, loss: 0.12261504679918289
step: 310, loss: 0.006869862787425518
step: 320, loss: 0.06610658764839172
step: 330, loss: 0.015379974618554115
step: 340, loss: 0.04359659552574158
step: 350, loss: 0.0010628130985423923
step: 360, loss: 0.08340112864971161
step: 370, loss: 0.18925416469573975
step: 380, loss: 0.028822094202041626
step: 390, loss: 0.12665413320064545
step: 400, loss: 0.05976913869380951
step: 410, loss: 0.049646731466054916
step: 420, loss: 0.012062128633260727
step: 430, loss: 0.014955226331949234
step: 440, loss: 0.055414535105228424
step: 450, loss: 0.022723300382494926
step: 460, loss: 0.032877158373594284
step: 470, loss: 0.039479807019233704
step: 480, loss: 0.2068672478199005
step: 490, loss: 0.01643744297325611
step: 500, loss: 0.1612488329410553
step: 510, loss: 0.015514673665165901
step: 520, loss: 0.04855310544371605
step: 530, loss: 0.00866367481648922
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9300373134328358, f1=0.9218604651162791,
best_f1=0.9218604651162791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014008976519107819
step: 10, loss: 0.014228101819753647
step: 20, loss: 0.11687395721673965
step: 30, loss: 0.017438432201743126

```

step: 40, loss: 0.018703745678067207
step: 50, loss: 0.05675927549600601
step: 60, loss: 0.013127898797392845
step: 70, loss: 0.002595438389107585
step: 80, loss: 0.001638023997657001
step: 90, loss: 0.016976656392216682
step: 100, loss: 0.07019287347793579
step: 110, loss: 0.004538485314697027
step: 120, loss: 0.00883584376424551
step: 130, loss: 0.004203883931040764
step: 140, loss: 0.0375799685716629
step: 150, loss: 0.06245533376932144
step: 160, loss: 0.03258798271417618
step: 170, loss: 0.1295512467622757
step: 180, loss: 0.022077038884162903
step: 190, loss: 0.001423089997842908
step: 200, loss: 0.013720179907977581
step: 210, loss: 0.07613172382116318
step: 220, loss: 0.05771968886256218
step: 230, loss: 0.09643890708684921
step: 240, loss: 0.0035572443157434464
step: 250, loss: 0.022076735273003578
step: 260, loss: 0.01774178259074688
step: 270, loss: 0.008652810007333755
step: 280, loss: 0.16737355291843414
step: 290, loss: 0.0019097456242889166
step: 300, loss: 0.020807037129998207
step: 310, loss: 0.008133841678500175
step: 320, loss: 0.02434352971613407
step: 330, loss: 0.002115862211212516
step: 340, loss: 0.0048689087852835655
step: 350, loss: 0.013193079270422459
step: 360, loss: 0.02832004986703396
step: 370, loss: 0.0042067524045705795
step: 380, loss: 0.01066708192229271
step: 390, loss: 0.004934691824018955
step: 400, loss: 0.009654135443270206
step: 410, loss: 0.0039640069007873535
step: 420, loss: 0.16462276875972748
step: 430, loss: 0.00950285978615284
step: 440, loss: 0.0021667887922376394
step: 450, loss: 0.08079519867897034
step: 460, loss: 0.012232815846800804
step: 470, loss: 0.013028601184487343
step: 480, loss: 0.01121152751147747
step: 490, loss: 0.005641681607812643
step: 500, loss: 0.02716829814016819
step: 510, loss: 0.0153029290959239


```

step: 520, loss: 0.18531550467014313
step: 530, loss: 0.10548549145460129
/usr/local/lib/python3.7/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=0.9296947271045328, f1=0.912882298424467,
best_f1=0.9218604651162791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03593999519944191
step: 10, loss: 0.015007133595645428
step: 20, loss: 0.0016690518241375685
step: 30, loss: 0.012942738831043243
step: 40, loss: 0.03770671784877777
step: 50, loss: 0.005144147202372551
step: 60, loss: 0.03470725938677788
step: 70, loss: 0.09062179177999496
step: 80, loss: 0.12140249460935593
step: 90, loss: 0.08369696140289307
step: 100, loss: 0.014008603058755398
step: 110, loss: 0.06227251887321472
step: 120, loss: 0.003128688782453537
step: 130, loss: 0.0013119817012920976
step: 140, loss: 0.004930784460157156
step: 150, loss: 0.013052236288785934
step: 160, loss: 0.0029952065087854862
step: 170, loss: 0.017514554783701897
step: 180, loss: 0.001692574005573988
step: 190, loss: 0.11332236230373383
step: 200, loss: 0.001557356328703463
step: 210, loss: 0.04353763163089752
step: 220, loss: 0.02754310704767704
step: 230, loss: 0.029778163880109787
step: 240, loss: 0.012866639532148838
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 250, loss: 0.02143000066280365
step: 260, loss: 0.0024012948852032423
step: 270, loss: 0.006164020858705044
step: 280, loss: 0.005797902587801218
step: 290, loss: 0.07706480473279953
step: 300, loss: 0.025053834542632103
step: 310, loss: 0.027207596227526665
step: 320, loss: 0.08889579027891159
step: 330, loss: 0.022207560017704964
step: 340, loss: 0.005076949018985033
step: 350, loss: 0.011877452954649925
step: 360, loss: 0.056753601878881454

```

```

step: 370, loss: 0.004360363353043795
step: 380, loss: 0.042703233659267426
step: 390, loss: 0.0017762985080480576
step: 400, loss: 0.002753408858552575
step: 410, loss: 0.03506613150238991
step: 420, loss: 0.004656198434531689
step: 430, loss: 0.0837777853012085
step: 440, loss: 0.0014668430667370558
step: 450, loss: 0.0013045119121670723
step: 460, loss: 0.004161639604717493
step: 470, loss: 0.0036744140088558197
step: 480, loss: 0.012732578441500664
step: 490, loss: 0.006430018227547407
step: 500, loss: 0.16251318156719208
step: 510, loss: 0.09729927778244019
step: 520, loss: 0.08309108763933182
step: 530, loss: 0.0035566387232393026
/usr/local/lib/python3.7/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=0.9232941176470588, f1=0.9155261915998112,
best_f1=0.9218604651162791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03275645151734352
step: 10, loss: 0.14850065112113953
step: 20, loss: 0.0034662929829210043
step: 30, loss: 0.0020458283834159374
step: 40, loss: 0.013448450714349747
step: 50, loss: 0.02759713865816593
step: 60, loss: 0.09527038037776947
step: 70, loss: 0.0066460915841162205
step: 80, loss: 0.005780573468655348
step: 90, loss: 0.0015320121310651302
step: 100, loss: 0.011220293119549751
step: 110, loss: 0.0004185680008959025
step: 120, loss: 0.04959959164261818
step: 130, loss: 0.001428509596735239
step: 140, loss: 0.0012334228958934546
step: 150, loss: 0.04435708746314049
step: 160, loss: 0.00043886867933906615
step: 170, loss: 0.021128058433532715
step: 180, loss: 0.0043291086331009865
step: 190, loss: 0.005187429022043943
step: 200, loss: 0.0008620452717877924
step: 210, loss: 0.009548197500407696
step: 220, loss: 0.0015005712630227208

```

```

step: 230, loss: 0.002317834412679076
step: 240, loss: 0.0016847428632900119
step: 250, loss: 0.0003422212030272931
step: 260, loss: 0.0012573704589158297
step: 270, loss: 0.00914551317691803
step: 280, loss: 0.007277489174157381
step: 290, loss: 0.15715013444423676
step: 300, loss: 0.0019153961911797523
step: 310, loss: 0.004313752055168152
step: 320, loss: 0.05333560332655907
step: 330, loss: 0.004124656319618225
step: 340, loss: 0.0035360995680093765
step: 350, loss: 0.004136071074754
step: 360, loss: 0.011964897625148296
step: 370, loss: 0.013600331731140614
step: 380, loss: 0.0012085438938811421
step: 390, loss: 0.00020390893041621894
step: 400, loss: 0.004129997454583645
step: 410, loss: 0.002939188852906227
step: 420, loss: 0.002509195590391755
step: 430, loss: 0.024034107103943825
step: 440, loss: 0.014702090062201023
step: 450, loss: 0.0019079884514212608
step: 460, loss: 0.005943169817328453
step: 470, loss: 0.004094367381185293
step: 480, loss: 0.08388858288526535
step: 490, loss: 0.0019085491076111794
step: 500, loss: 0.004026051610708237
step: 510, loss: 0.11381292343139648
step: 520, loss: 0.0006899121799506247
step: 530, loss: 0.003952185623347759
/usr/local/lib/python3.7/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=0.9112709832134291, f1=0.9012464046021093,
best_f1=0.9218604651162791
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009717652574181557
step: 10, loss: 0.0006255880580283701
step: 20, loss: 0.09690696746110916
step: 30, loss: 0.0005412728060036898
step: 40, loss: 0.006090562324970961
step: 50, loss: 0.008341302163898945
step: 60, loss: 0.0007399183814413846
step: 70, loss: 0.0010859956964850426
step: 80, loss: 0.0011331363348290324

```

```
step: 90, loss: 0.000214208266697824
step: 100, loss: 0.0012064214097335935
step: 110, loss: 0.0002619070000946522
step: 120, loss: 0.0022474525030702353
step: 130, loss: 0.00027654547011479735
step: 140, loss: 0.002372988034039736
step: 150, loss: 0.01943478174507618
step: 160, loss: 0.004082463216036558
step: 170, loss: 0.0001357924338662997
step: 180, loss: 0.003515472635626793
step: 190, loss: 0.02279866859316826
step: 200, loss: 0.0005285681108944118
step: 210, loss: 0.13318997621536255
step: 220, loss: 0.0027345577254891396
step: 230, loss: 0.012179802171885967
step: 240, loss: 0.030613616108894348
step: 250, loss: 0.003968267235904932
step: 260, loss: 0.0006256437627598643
step: 270, loss: 0.024270372465252876
step: 280, loss: 0.0015473227249458432
step: 290, loss: 0.00016761515871621668
step: 300, loss: 0.0031451350077986717
step: 310, loss: 0.007111575920134783
step: 320, loss: 0.0009512068936601281
step: 330, loss: 0.004690965637564659
step: 340, loss: 0.08184810727834702
step: 350, loss: 0.00113730039447546
step: 360, loss: 0.04919832572340965
step: 370, loss: 0.0028224620036780834
step: 380, loss: 0.00034310147748328745
step: 390, loss: 0.01899496465921402
step: 400, loss: 0.002446339698508382
step: 410, loss: 0.006222357973456383
step: 420, loss: 0.007608650717884302
step: 430, loss: 0.0026178492698818445
step: 440, loss: 8.382948726648465e-05
step: 450, loss: 0.0001219103651237674
step: 460, loss: 7.160489622037858e-05
step: 470, loss: 0.0006084393244236708
step: 480, loss: 0.070460245013237
step: 490, loss: 0.013174567371606827
step: 500, loss: 0.003753578057512641
step: 510, loss: 0.002600171137601137
step: 520, loss: 0.0010049634147435427
step: 530, loss: 0.1311797797679901
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9348623853211009, f1=0.9278256922378576,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029702650499530137
step: 10, loss: 0.0031468886882066727
step: 20, loss: 0.00042638974264264107
step: 30, loss: 0.00805203802883625
step: 40, loss: 0.019255202263593674
step: 50, loss: 0.00030918308766558766
step: 60, loss: 0.007502499967813492
step: 70, loss: 0.000845936534460634
step: 80, loss: 0.0009930958040058613
step: 90, loss: 0.002477996051311493
step: 100, loss: 7.53315252950415e-05
step: 110, loss: 0.00026377104222774506
step: 120, loss: 0.0006676369230262935
step: 130, loss: 0.011751458048820496
step: 140, loss: 6.240420771064237e-05
step: 150, loss: 0.0005458152736537158
step: 160, loss: 0.0004928645212203264
step: 170, loss: 0.0009287514840252697
step: 180, loss: 0.0006313444464467466
step: 190, loss: 0.0008152038208208978
step: 200, loss: 8.769274427322671e-05
step: 210, loss: 0.00010922650835709646
step: 220, loss: 7.711737271165475e-05
step: 230, loss: 0.09495517611503601
step: 240, loss: 0.0008268069941550493
step: 250, loss: 0.0005990591598674655
step: 260, loss: 0.00026893592439591885
step: 270, loss: 0.0004867773677688092
step: 280, loss: 0.00029208199703134596
step: 290, loss: 0.0008317812462337315
step: 300, loss: 0.018860487267374992
step: 310, loss: 0.00017523285350762308
step: 320, loss: 0.0002883501583710313
step: 330, loss: 0.001998076681047678
step: 340, loss: 0.0036317245103418827
step: 350, loss: 0.0007358689908869565
step: 360, loss: 0.002224421827122569
step: 370, loss: 0.0006291468744166195
step: 380, loss: 0.0012155374279245734

```

```
step: 390, loss: 0.001196654629893601
step: 400, loss: 0.003173787146806717
step: 410, loss: 0.0021389981266111135
step: 420, loss: 0.006606516428291798
step: 430, loss: 0.00017134027439169586
step: 440, loss: 0.0007808374939486384
step: 450, loss: 0.0009281456004828215
step: 460, loss: 0.0007187701994553208
step: 470, loss: 0.0021278050262480974
step: 480, loss: 0.12227635830640793
step: 490, loss: 0.0009297503856942058
step: 500, loss: 0.0003088205121457577
step: 510, loss: 0.0011131507344543934
step: 520, loss: 0.0007617485825903714
step: 530, loss: 0.02731630951166153
/usr/local/lib/python3.7/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=0.9252767527675276, f1=0.9227230910763569,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021567470394074917
step: 10, loss: 0.0325743593275547
step: 20, loss: 0.0003908572834916413
step: 30, loss: 0.12153588980436325
step: 40, loss: 0.00010392047261120752
step: 50, loss: 0.006427675951272249
step: 60, loss: 0.04063229262828827
step: 70, loss: 0.0003957074659410864
step: 80, loss: 0.014473183080554008
step: 90, loss: 0.0009001547005027533
step: 100, loss: 0.00038270457298494875
step: 110, loss: 0.00022471662668976933
step: 120, loss: 0.0003132352139800787
step: 130, loss: 0.00012956908904016018
step: 140, loss: 0.004699037875980139
step: 150, loss: 0.010922005400061607
step: 160, loss: 0.0314205177128315
step: 170, loss: 0.0004045121022500098
step: 180, loss: 0.0001597420050529763
step: 190, loss: 0.0017349711852148175
step: 200, loss: 0.0011247514048591256
step: 210, loss: 0.002519201021641493
step: 220, loss: 0.002089668298140168
step: 230, loss: 0.00014393069432117045
step: 240, loss: 0.002849096432328224
```

```

step: 250, loss: 0.0010960143990814686
step: 260, loss: 0.0011174783576279879
step: 270, loss: 0.18505483865737915
step: 280, loss: 0.09660976380109787
step: 290, loss: 0.008944257162511349
step: 300, loss: 0.0007000002660788596
step: 310, loss: 0.0015785802388563752
step: 320, loss: 0.006594059988856316
step: 330, loss: 0.00031924102222546935
step: 340, loss: 0.0035907926503568888
step: 350, loss: 0.00020890971063636243
step: 360, loss: 0.00015910038200672716
step: 370, loss: 0.02411090023815632
step: 380, loss: 0.07510209828615189
step: 390, loss: 0.10067986696958542
step: 400, loss: 0.0010486847022548318
step: 410, loss: 0.00010114223550772294
step: 420, loss: 0.0019670359324663877
step: 430, loss: 0.03044169209897518
step: 440, loss: 0.029670586809515953
step: 450, loss: 0.005527024623006582
step: 460, loss: 4.961053855367936e-05
step: 470, loss: 0.0003416603140067309
step: 480, loss: 7.052377623040229e-05
step: 490, loss: 0.0003253297763876617
step: 500, loss: 0.0004679708508774638
step: 510, loss: 0.004676567390561104
step: 520, loss: 0.00397426076233387
step: 530, loss: 3.443885361775756e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9237918215613383, f1=0.9068150208623087,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010518726194277406
step: 10, loss: 0.0004215231165289879
step: 20, loss: 0.0011014258489012718
step: 30, loss: 0.00020159172709099948
step: 40, loss: 4.903268563793972e-05
step: 50, loss: 7.94461157056503e-05
step: 60, loss: 3.9262584323296323e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.003044678131118417
step: 80, loss: 0.00028899224707856774
step: 90, loss: 8.15197272459045e-05

```

```
step: 100, loss: 3.188003756804392e-05
step: 110, loss: 0.0006740832468494773
step: 120, loss: 0.00037787805194966495
step: 130, loss: 9.256205521523952e-05
step: 140, loss: 3.673361425171606e-05
step: 150, loss: 6.111735274316743e-05
step: 160, loss: 0.002722078002989292
step: 170, loss: 0.005953626707196236
step: 180, loss: 2.8165943149360828e-05
step: 190, loss: 4.217287278152071e-05
step: 200, loss: 0.00018328483565710485
step: 210, loss: 3.6404599086381495e-05
step: 220, loss: 0.0028994856402277946
step: 230, loss: 4.230867853038944e-05
step: 240, loss: 8.259883179562166e-05
step: 250, loss: 6.376235978677869e-05
step: 260, loss: 0.04552552476525307
step: 270, loss: 2.204218617407605e-05
step: 280, loss: 3.338500391691923e-05
step: 290, loss: 0.058656591922044754
step: 300, loss: 3.866892075166106e-05
step: 310, loss: 0.012051098980009556
step: 320, loss: 0.03243112564086914
step: 330, loss: 0.0002309717092430219
step: 340, loss: 6.27697809250094e-05
step: 350, loss: 0.0027693151496350765
step: 360, loss: 2.7611231416813098e-05
step: 370, loss: 0.00047974183689802885
step: 380, loss: 0.0003628809645306319
step: 390, loss: 0.003485799068585038
step: 400, loss: 0.0010947661940008402
step: 410, loss: 0.0029835153836756945
step: 420, loss: 6.733314512530342e-05
step: 430, loss: 0.0023202192969620228
step: 440, loss: 0.0044007194228470325
step: 450, loss: 0.0002662469632923603
step: 460, loss: 0.00016189305461011827
step: 470, loss: 0.00015188861289061606
step: 480, loss: 5.715875158784911e-05
step: 490, loss: 0.0013444741489365697
step: 500, loss: 0.0015115300193428993
step: 510, loss: 0.016797633841633797
step: 520, loss: 0.0011844616383314133
step: 530, loss: 5.134392631589435e-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=0.9325210871602625, f1=0.9282700421940927,
```



```
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.953997894423082e-05
step: 10, loss: 5.516310557140969e-05
step: 20, loss: 2.7067086193710566e-05
step: 30, loss: 4.673732109949924e-05
step: 40, loss: 6.644726090598851e-05
step: 50, loss: 9.025391045724973e-05
step: 60, loss: 7.321356679312885e-05
step: 70, loss: 0.0066113220527768135
step: 80, loss: 5.1788578275591135e-05
step: 90, loss: 0.00014795952301938087
step: 100, loss: 0.0005385371041484177
step: 110, loss: 2.6314826754969545e-05
step: 120, loss: 0.037015125155448914
step: 130, loss: 3.522102633723989e-05
step: 140, loss: 0.00011627024650806561
step: 150, loss: 5.334442175808363e-05
step: 160, loss: 5.938001049798913e-05
step: 170, loss: 4.180668111075647e-05
step: 180, loss: 4.136320058023557e-05
step: 190, loss: 0.00010666754678823054
step: 200, loss: 0.0031523555517196655
step: 210, loss: 0.00018304250261280686
step: 220, loss: 0.00033248536055907607
step: 230, loss: 0.0035157182719558477
step: 240, loss: 7.717379776295274e-05
step: 250, loss: 4.5187800424173474e-05
step: 260, loss: 0.0019404874183237553
step: 270, loss: 0.000981162185780704
step: 280, loss: 0.00012357922969385982
step: 290, loss: 0.00019831440295092762
step: 300, loss: 2.401644815108739e-05
step: 310, loss: 3.506418943288736e-05
step: 320, loss: 0.0014711221447214484
step: 330, loss: 0.00010687114991014823
step: 340, loss: 0.0001610025210538879
step: 350, loss: 0.00020606584439519793
step: 360, loss: 0.0006825198652222753
step: 370, loss: 0.0003624722012318671
step: 380, loss: 0.0008317400352098048
step: 390, loss: 0.0002059179823845625
step: 400, loss: 0.06050065532326698
step: 410, loss: 0.0005206117639318109
step: 420, loss: 0.20837639272212982
step: 430, loss: 3.551000190782361e-05
```

```

step: 440, loss: 0.006998415570706129
step: 450, loss: 6.10796341788955e-05
step: 460, loss: 0.0017823355738073587
step: 470, loss: 0.00019142328528687358
step: 480, loss: 5.7366218243259937e-05
step: 490, loss: 0.0012123020133003592
step: 500, loss: 0.0037286588922142982
step: 510, loss: 2.9536851798184216e-05
step: 520, loss: 0.00010933510202448815
step: 530, loss: 0.00042793998727574944
/usr/local/lib/python3.7/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=0.930276087973795, f1=0.9185882352941176,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019798825087491423
step: 10, loss: 0.000497575500048697
step: 20, loss: 0.00011290224210824817
step: 30, loss: 5.0497612392064184e-05
step: 40, loss: 8.503944263793528e-05
step: 50, loss: 0.00013377676077652723
step: 60, loss: 0.038123685866594315
step: 70, loss: 2.4213823053287342e-05
step: 80, loss: 3.296724389656447e-05
step: 90, loss: 3.4430588129907846e-05
step: 100, loss: 0.0106945326551795
step: 110, loss: 8.502614946337417e-05
step: 120, loss: 0.00016156914352905005
step: 130, loss: 0.0019384657498449087
step: 140, loss: 0.0005384988035075366
step: 150, loss: 1.3980839867144823e-05
step: 160, loss: 0.00014076569641474634
step: 170, loss: 2.908230817411095e-05
step: 180, loss: 5.073121064924635e-05
step: 190, loss: 0.0028150649741292
step: 200, loss: 0.006271859630942345
step: 210, loss: 0.0001491977454861626
step: 220, loss: 0.0009648512350395322
step: 230, loss: 7.799485319992527e-05
step: 240, loss: 0.00039094293606467545
step: 250, loss: 2.1591253243968822e-05
step: 260, loss: 2.1911677322350442e-05
step: 270, loss: 0.0001110882731154561
step: 280, loss: 9.582141501596197e-05
step: 290, loss: 2.0753101125592366e-05

```

```

step: 300, loss: 8.176191477105021e-05
step: 310, loss: 0.00020126021991018206
step: 320, loss: 0.00010498863412067294
step: 330, loss: 0.0003830725618172437
step: 340, loss: 0.0004551328020170331
step: 350, loss: 0.0003259831282775849
step: 360, loss: 9.633307490730658e-05
step: 370, loss: 0.012995137833058834
step: 380, loss: 0.0004256577230989933
step: 390, loss: 0.0002032285847235471
step: 400, loss: 0.0002917663659900427
step: 410, loss: 5.5115062423283234e-05
step: 420, loss: 0.0030569378286600113
step: 430, loss: 3.6503806768450886e-05
step: 440, loss: 0.0005455637583509088
step: 450, loss: 0.00020357989706099033
step: 460, loss: 0.00013785772898700088
step: 470, loss: 6.215825851541013e-05
step: 480, loss: 0.0009581110207363963
step: 490, loss: 5.574045280809514e-05
step: 500, loss: 0.002352561568841338
step: 510, loss: 0.0003878124407492578
step: 520, loss: 4.4074673496652395e-05
step: 530, loss: 0.00040188629645854235
/usr/local/lib/python3.7/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=0.9295904279797514, f1=0.9220055710306406,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017259805463254452
step: 10, loss: 2.2001144316163845e-05
step: 20, loss: 0.00012793521455023438
step: 30, loss: 5.604374746326357e-05
step: 40, loss: 0.0001089470533770509
step: 50, loss: 0.019839158281683922
step: 60, loss: 0.002913211937993765
step: 70, loss: 7.374497363343835e-05
step: 80, loss: 0.0003047978680115193
step: 90, loss: 5.3755156841361895e-05
step: 100, loss: 0.00015990981773938984
step: 110, loss: 0.0016918821493163705
step: 120, loss: 1.5388974134111777e-05
step: 130, loss: 0.00010909508273471147
step: 140, loss: 2.6354833607911132e-05
step: 150, loss: 3.578404721338302e-05

```

```

step: 160, loss: 7.0178386522457e-05
step: 170, loss: 0.013829309493303299
step: 180, loss: 0.00025598282809369266
step: 190, loss: 5.471969780046493e-05
step: 200, loss: 3.2069827284431085e-05
step: 210, loss: 5.825484186061658e-05
step: 220, loss: 7.138562068575993e-05
step: 230, loss: 1.5735397028038278e-05
step: 240, loss: 0.0005922638811171055
step: 250, loss: 2.0991594283259474e-05
step: 260, loss: 1.742294625728391e-05
step: 270, loss: 6.011364530422725e-05
step: 280, loss: 0.0007435836014337838
step: 290, loss: 2.7953630706178956e-05
step: 300, loss: 0.0004296534461900592
step: 310, loss: 2.45488863583887e-05
step: 320, loss: 1.8629856640473008e-05
step: 330, loss: 7.879081385908648e-05
step: 340, loss: 3.287247091066092e-05
step: 350, loss: 3.950058453483507e-05
step: 360, loss: 5.872672772966325e-05
step: 370, loss: 0.0005725902155973017
step: 380, loss: 2.9074790290906094e-05
step: 390, loss: 2.7711468646884896e-05
step: 400, loss: 0.0001516190532129258
step: 410, loss: 0.00019252007768955082
step: 420, loss: 0.0011221329914405942
step: 430, loss: 2.1989924789522775e-05
step: 440, loss: 3.642441879492253e-05
step: 450, loss: 0.0023591527715325356
step: 460, loss: 1.3813236364512704e-05
step: 470, loss: 1.7631484297453426e-05
step: 480, loss: 0.06006309762597084
step: 490, loss: 0.00037615405744872987
step: 500, loss: 0.0007923663943074644
step: 510, loss: 0.0006404114537872374
step: 520, loss: 0.00014365841343533248
step: 530, loss: 3.097115404671058e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9306569343065694, f1=0.91919191919192,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038992136251181364
step: 10, loss: 6.575549195986241e-05

```

step: 20, loss: 0.00032533545163460076
step: 30, loss: 9.603805665392429e-05
step: 40, loss: 1.910670289362315e-05
step: 50, loss: 0.0004989627050235868
step: 60, loss: 0.0001712052326183766
step: 70, loss: 3.462704989942722e-05
step: 80, loss: 0.00011092858039774
step: 90, loss: 0.0003025145852006972
step: 100, loss: 0.0006104277563281357
step: 110, loss: 5.740196502301842e-05
step: 120, loss: 0.004268335644155741
step: 130, loss: 2.2917369278729893e-05
step: 140, loss: 8.703325875103474e-05
step: 150, loss: 0.0003484599874354899
step: 160, loss: 3.3379008527845144e-05
step: 170, loss: 0.00031466654036194086
step: 180, loss: 3.3938922570087016e-05
step: 190, loss: 0.0012479603756219149
step: 200, loss: 7.65460790717043e-05
step: 210, loss: 4.048154369229451e-05
step: 220, loss: 1.4580617971660104e-05
step: 230, loss: 0.0010647507151588798
step: 240, loss: 0.00012575041910167783
step: 250, loss: 1.4860031114949379e-05
step: 260, loss: 1.548584805277642e-05
step: 270, loss: 1.381693255098071e-05
step: 280, loss: 0.0002099053090205416
step: 290, loss: 0.0004025642410852015
step: 300, loss: 1.8443633962306194e-05
step: 310, loss: 0.0021895423997193575
step: 320, loss: 0.0067748394794762135
step: 330, loss: 0.0015096277929842472
step: 340, loss: 2.9665487090824172e-05
step: 350, loss: 3.0149181839078665e-05
step: 360, loss: 4.2676641896832734e-05
step: 370, loss: 0.00015289433940779418
step: 380, loss: 1.687530311755836e-05
step: 390, loss: 0.00011453829938545823
step: 400, loss: 9.149884135695174e-05
step: 410, loss: 0.0003411297802813351
step: 420, loss: 1.6867836166056804e-05
step: 430, loss: 7.16036229277961e-05
step: 440, loss: 5.916254667681642e-05
step: 450, loss: 9.083474287763238e-05
step: 460, loss: 0.0006909787771292031
step: 470, loss: 0.00014346376701723784
step: 480, loss: 9.374086221214384e-05
step: 490, loss: 0.00011633334361249581

```

step: 500, loss: 1.1473794984340202e-05
step: 510, loss: 3.402333095436916e-05
step: 520, loss: 0.004052670672535896
step: 530, loss: 1.3094268979330081e-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=0.9197707736389685, f1=0.9108433734939759,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002784297685138881
step: 10, loss: 1.1082654964411631e-05
step: 20, loss: 1.6726133253541775e-05
step: 30, loss: 3.476894198684022e-05
step: 40, loss: 2.727501305344049e-05
step: 50, loss: 4.884713780484162e-05
step: 60, loss: 1.1533405995578505e-05
step: 70, loss: 0.004368986003100872
step: 80, loss: 2.946895983768627e-05
step: 90, loss: 1.878251350717619e-05
step: 100, loss: 0.0003913629916496575
step: 110, loss: 2.2626658392255194e-05
step: 120, loss: 1.5139382412598934e-05
step: 130, loss: 0.00019934881129302084
step: 140, loss: 3.9127011405071244e-05
step: 150, loss: 1.27180001072702e-05
step: 160, loss: 8.767870895098895e-05
step: 170, loss: 1.9050681657972746e-05
step: 180, loss: 2.8138227207819e-05
step: 190, loss: 6.162712816148996e-05
step: 200, loss: 0.008522801101207733
step: 210, loss: 1.5564066416118294e-05
step: 220, loss: 1.543366306577809e-05
step: 230, loss: 3.0277384212240577e-05
step: 240, loss: 3.1778556149220094e-05
step: 250, loss: 1.8387738236924633e-05
step: 260, loss: 0.00011496499791974202
step: 270, loss: 7.537949568359181e-05
step: 280, loss: 1.4979223124100827e-05
step: 290, loss: 5.538707773666829e-05
step: 300, loss: 3.445563925197348e-05
step: 310, loss: 1.912532388814725e-05
step: 320, loss: 0.001325904275290668
step: 330, loss: 3.885431942762807e-05
step: 340, loss: 2.2224501663004048e-05
step: 350, loss: 2.4220671548391692e-05

```

```

step: 360, loss: 0.0016949607525020838
step: 370, loss: 4.504420212469995e-05
step: 380, loss: 0.00039932687650434673
step: 390, loss: 0.01378073263913393
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 0.00028342296718619764
step: 410, loss: 2.2823789549875073e-05
step: 420, loss: 0.0002697133750189096
step: 430, loss: 2.539728484407533e-05
step: 440, loss: 3.030016341654118e-05
step: 450, loss: 0.0002515902160666883
step: 460, loss: 4.651767449104227e-05
step: 470, loss: 0.0010555799817666411
step: 480, loss: 1.6841791875776835e-05
step: 490, loss: 2.1594481950160116e-05
step: 500, loss: 1.3328954082680866e-05
step: 510, loss: 0.0007344963378272951
step: 520, loss: 7.102434028638527e-05
step: 530, loss: 4.8899197281571105e-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=0.9296413980935089, f1=0.920562868815252,
best_f1=0.9278256922378576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.5230412575183436e-05
step: 10, loss: 2.6357338356319815e-05
step: 20, loss: 9.680917719379067e-05
step: 30, loss: 0.004280685912817717
step: 40, loss: 0.07986417412757874
step: 50, loss: 1.726274422253482e-05
step: 60, loss: 2.4352761101908982e-05
step: 70, loss: 0.0007425005314871669
step: 80, loss: 1.9777127818088047e-05
step: 90, loss: 1.9654218704090454e-05
step: 100, loss: 6.840448622824624e-05
step: 110, loss: 6.459037831518799e-05
step: 120, loss: 0.005219500977545977
step: 130, loss: 0.0021991031244397163
step: 140, loss: 1.6186158973141573e-05
step: 150, loss: 1.6260695701930672e-05
step: 160, loss: 5.457469160319306e-05
step: 170, loss: 3.5270819353172556e-05
step: 180, loss: 1.3444415344565641e-05
step: 190, loss: 1.7091433619498275e-05
step: 200, loss: 3.832177753793076e-05

```

```

step: 210, loss: 9.147867240244523e-05
step: 220, loss: 1.287448230868904e-05
step: 230, loss: 1.4480055142485071e-05
step: 240, loss: 4.3180392822250724e-05
step: 250, loss: 2.169928848161362e-05
step: 260, loss: 0.00020952941849827766
step: 270, loss: 1.3656783266924322e-05
step: 280, loss: 2.241042056994047e-05
step: 290, loss: 1.4774324881727807e-05
step: 300, loss: 1.6998121282085776e-05
step: 310, loss: 3.171491334796883e-05
step: 320, loss: 5.151388904778287e-05
step: 330, loss: 2.094288902299013e-05
step: 340, loss: 2.043182212219108e-05
step: 350, loss: 1.7902815670822747e-05
step: 360, loss: 0.0019007368246093392
step: 370, loss: 2.3533852072432637e-05
step: 380, loss: 1.4815322174399626e-05
step: 390, loss: 2.4179929823731072e-05
step: 400, loss: 1.8152963093598373e-05
step: 410, loss: 1.733340286591556e-05
step: 420, loss: 2.8999387723160908e-05
step: 430, loss: 3.46120104950387e-05
step: 440, loss: 2.802681592584122e-05
step: 450, loss: 9.258893987862393e-05
step: 460, loss: 2.083113940898329e-05
step: 470, loss: 0.0013442394556477666
step: 480, loss: 1.1365776117600035e-05
step: 490, loss: 1.2766455256496556e-05
step: 500, loss: 1.8488050045561977e-05
step: 510, loss: 5.6456003221683204e-05
step: 520, loss: 1.3247007700556424e-05
step: 530, loss: 1.6260657503153197e-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=0.9291994447015272, f1=0.919944470152707,
best_f1=0.9278256922378576
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 [00:29, 192.83it/s]
load_f1 = 0.9348025711662074
real_f1 = 0.9337626494940202
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, 189.85it/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 bert-base-uncased \
    --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 bert-base-uncased \
  --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: 0.5425687432289124
step: 10, loss: 0.36731481552124023
step: 20, loss: 0.36463871598243713
step: 30, loss: 0.3450835049152374
step: 40, loss: 0.1637304127216339
step: 50, loss: 0.4170503616333008
step: 60, loss: 0.2843277156352997
step: 70, loss: 0.18189765512943268
step: 80, loss: 0.18742412328720093
step: 90, loss: 0.3964230418205261
step: 100, loss: 0.34028610587120056
step: 110, loss: 0.21414773166179657
step: 120, loss: 0.20328186452388763
step: 130, loss: 0.1896314173936844
step: 140, loss: 0.19417986273765564
step: 150, loss: 0.18473398685455322
step: 160, loss: 0.34342649579048157
step: 170, loss: 0.2860686779022217
step: 180, loss: 0.14511999487876892
step: 190, loss: 0.2608819901943207
step: 200, loss: 0.2991541624069214
step: 210, loss: 0.25575709342956543
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.564755838641189, f1=0.6171171171171171,
best_f1=0.6171171171171171
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19796344637870789
step: 10, loss: 0.23283670842647552
step: 20, loss: 0.1668589860200882
step: 30, loss: 0.14008890092372894
step: 40, loss: 0.2721114754676819
step: 50, loss: 0.07193109393119812
step: 60, loss: 0.3217650055885315
step: 70, loss: 0.09714290499687195
step: 80, loss: 0.254610538482666
step: 90, loss: 0.10286614298820496
step: 100, loss: 0.009485863149166107
step: 110, loss: 0.1453656554222107
step: 120, loss: 0.1422993689775467

```

```

step: 130, loss: 0.06980980187654495
step: 140, loss: 0.3103395104408264
step: 150, loss: 0.2598874270915985
step: 160, loss: 0.20402014255523682
step: 170, loss: 0.09425190091133118
step: 180, loss: 0.24440698325634003
step: 190, loss: 0.20936639606952667
step: 200, loss: 0.04158642515540123
step: 210, loss: 0.15540482103824615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.605427974947808, f1=0.6311300639658849,
best_f1=0.6311300639658849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.049943145364522934
step: 10, loss: 0.10010017454624176
step: 20, loss: 0.189909428358078
step: 30, loss: 0.18689729273319244
step: 40, loss: 0.07045457512140274
step: 50, loss: 0.08209437876939774
step: 60, loss: 0.2106439173221588
step: 70, loss: 0.05555292218923569
step: 80, loss: 0.2028937041759491
step: 90, loss: 0.05825386568903923
step: 100, loss: 0.13041092455387115
step: 110, loss: 0.16435886919498444
step: 120, loss: 0.1538999378681183
step: 130, loss: 0.2184806615114212
step: 140, loss: 0.130941703915596
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.19806480407714844
step: 160, loss: 0.012422798201441765
step: 170, loss: 0.12907518446445465
step: 180, loss: 0.06325237452983856
step: 190, loss: 0.22762931883335114
step: 200, loss: 0.049790892750024796
step: 210, loss: 0.24946291744709015
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.6197183098591549, f1=0.6452905811623246,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08901810646057129
step: 10, loss: 0.05527091026306152
step: 20, loss: 0.05913827195763588
step: 30, loss: 0.1267441213130951
step: 40, loss: 0.02113049291074276
step: 50, loss: 0.20276795327663422
step: 60, loss: 0.07627850025892258
step: 70, loss: 0.23380520939826965
step: 80, loss: 0.03699542582035065
step: 90, loss: 0.008605497889220715
step: 100, loss: 0.17128701508045197
step: 110, loss: 0.05548013746738434
step: 120, loss: 0.06757694482803345
step: 130, loss: 0.16498441994190216
step: 140, loss: 0.25511959195137024
step: 150, loss: 0.030442196875810623
step: 160, loss: 0.04993711784482002
step: 170, loss: 0.09241853654384613
step: 180, loss: 0.25081866979599
step: 190, loss: 0.04420215263962746
step: 200, loss: 0.12231796979904175
step: 210, loss: 0.18145586550235748
/usr/local/lib/python3.7/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=0.6191446028513239, f1=0.6398305084745763,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16056528687477112
step: 10, loss: 0.03527902811765671
step: 20, loss: 0.05844531208276749
step: 30, loss: 0.0321698933839798
step: 40, loss: 0.03394206985831261
step: 50, loss: 0.087093286216259
step: 60, loss: 0.1676856428384781
step: 70, loss: 0.06808336824178696
step: 80, loss: 0.12461848556995392
step: 90, loss: 0.18976108729839325
```

```

step: 100, loss: 0.004807612858712673
step: 110, loss: 0.16880223155021667
step: 120, loss: 0.053899724036455154
step: 130, loss: 0.07762877643108368
step: 140, loss: 0.03530817851424217
step: 150, loss: 0.02911079302430153
step: 160, loss: 0.13215157389640808
step: 170, loss: 0.05948077514767647
step: 180, loss: 0.11514181643724442
step: 190, loss: 0.015856973826885223
step: 200, loss: 0.2680121660232544
step: 210, loss: 0.03144887089729309
/usr/local/lib/python3.7/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=0.5754527162977867, f1=0.6198347107438017,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021839873865246773
step: 10, loss: 0.021856199949979782
step: 20, loss: 0.003997281193733215
step: 30, loss: 0.0007676546229049563
step: 40, loss: 0.0465519018471241
step: 50, loss: 0.05598337948322296
step: 60, loss: 0.019257213920354843
step: 70, loss: 0.015618446283042431
step: 80, loss: 0.24137574434280396
step: 90, loss: 0.07017102092504501
step: 100, loss: 0.0424807146191597
step: 110, loss: 0.0064195310696959496
step: 120, loss: 0.08615506440401077
step: 130, loss: 0.07364015281200409
step: 140, loss: 0.007018789649009705
step: 150, loss: 0.010599980130791664
step: 160, loss: 0.011315410025417805
step: 170, loss: 0.11321219801902771
step: 180, loss: 0.027111347764730453
step: 190, loss: 0.04408985748887062
step: 200, loss: 0.01989610306918621
step: 210, loss: 0.038333695381879807
/usr/local/lib/python3.7/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=0.5973025048169557, f1=0.6215139442231076,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019642503932118416
step: 10, loss: 0.009694666601717472
step: 20, loss: 0.016831567510962486
step: 30, loss: 0.010769758373498917
step: 40, loss: 0.09108764678239822
step: 50, loss: 0.05035245418548584
step: 60, loss: 0.016459660604596138
step: 70, loss: 0.007867238484323025
step: 80, loss: 0.029674984514713287
step: 90, loss: 0.09884826093912125
step: 100, loss: 0.009474498219788074
step: 110, loss: 0.04288423806428909
step: 120, loss: 0.06136512756347656
step: 130, loss: 0.015496695414185524
step: 140, loss: 0.037813130766153336
step: 150, loss: 0.0061620245687663555
step: 160, loss: 0.02529778704047203
step: 170, loss: 0.0016564549878239632
step: 180, loss: 0.013467421755194664
step: 190, loss: 0.022861571982502937
step: 200, loss: 0.030961347743868828
step: 210, loss: 0.0035480125807225704
/usr/local/lib/python3.7/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=0.5545851528384279, f1=0.6133333333333332,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03950076177716255
step: 10, loss: 0.02443673647940159
step: 20, loss: 0.01397599745541811
step: 30, loss: 0.0026972410269081593
step: 40, loss: 0.006648447830229998
step: 50, loss: 0.004266645293682814
step: 60, loss: 0.13876011967658997
step: 70, loss: 0.014708274975419044
step: 80, loss: 0.015690673142671585
step: 90, loss: 0.02340044640004635
step: 100, loss: 0.03151620551943779
step: 110, loss: 0.02343909814953804
step: 120, loss: 0.0039958711713552475
step: 130, loss: 0.0014359727501869202
step: 140, loss: 0.013118883594870567
step: 150, loss: 0.008754375390708447

```

```

step: 160, loss: 0.035399485379457474
step: 170, loss: 0.016473809257149696
step: 180, loss: 0.09233017265796661
step: 190, loss: 0.0025134612806141376
step: 200, loss: 0.03951618820428848
step: 210, loss: 0.08708585798740387
/usr/local/lib/python3.7/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=0.5653104925053533, f1=0.6274509803921569,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014830072410404682
step: 10, loss: 0.008112623356282711
step: 20, loss: 0.020197054371237755
step: 30, loss: 0.015750406309962273
step: 40, loss: 0.02512260526418686
step: 50, loss: 0.0004133554466534406
step: 60, loss: 0.01461930200457573
step: 70, loss: 0.0008544484153389931
step: 80, loss: 0.002765498124063015
step: 90, loss: 0.0014442631509155035
step: 100, loss: 0.0006953014526516199
step: 110, loss: 0.06084620952606201
step: 120, loss: 0.015831809490919113
step: 130, loss: 0.0007693234947510064
step: 140, loss: 0.0022031632252037525
step: 150, loss: 0.03530101105570793
step: 160, loss: 0.0003288098960183561
step: 170, loss: 0.0028187076095491648
step: 180, loss: 0.04494944214820862
step: 190, loss: 0.0004155482747592032
step: 200, loss: 0.03662963956594467
step: 210, loss: 0.0022158699575811625
/usr/local/lib/python3.7/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=0.54337899543379, f1=0.5967365967365967,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010666750371456146
step: 10, loss: 0.0038399246986955404
step: 20, loss: 0.0018683267990127206
step: 30, loss: 0.005783209577202797

```



```

step: 40, loss: 0.0039078062400221825
step: 50, loss: 0.0010473368456587195
step: 60, loss: 0.047945279628038406
step: 70, loss: 0.022387400269508362
step: 80, loss: 0.010334371589124203
step: 90, loss: 0.042907603085041046
step: 100, loss: 0.01668045111000538
step: 110, loss: 0.0005535919335670769
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 120, loss: 0.0029512846376746893
step: 130, loss: 0.000823818554636091
step: 140, loss: 0.004086012486368418
step: 150, loss: 0.08183100074529648
step: 160, loss: 0.0031296999659389257
step: 170, loss: 0.0009000185527838767
step: 180, loss: 0.1677587330341339
step: 190, loss: 0.06168324872851372
step: 200, loss: 0.0010898024775087833
step: 210, loss: 0.05065380409359932
/usr/local/lib/python3.7/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=0.5688888888888889, f1=0.6031746031746031,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011139529757201672
step: 10, loss: 0.0035929526202380657
step: 20, loss: 0.007086622528731823
step: 30, loss: 0.009360265918076038
step: 40, loss: 0.02151186764240265
step: 50, loss: 0.0010682233842089772
step: 60, loss: 0.016486940905451775
step: 70, loss: 0.005988935008645058
step: 80, loss: 0.07821740210056305
step: 90, loss: 0.039708394557237625
step: 100, loss: 0.011555788107216358
step: 110, loss: 0.02382788248360157
step: 120, loss: 0.05630321800708771
step: 130, loss: 0.021465683355927467
step: 140, loss: 0.06567580997943878
step: 150, loss: 0.00013107484846841544
step: 160, loss: 0.0009294373448938131
step: 170, loss: 0.022342130541801453
step: 180, loss: 0.0004313152749091387
step: 190, loss: 0.00025951897259801626
step: 200, loss: 0.0005623206961899996

```

```

step: 210, loss: 0.0001653011131566018
/usr/local/lib/python3.7/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=0.5662921348314606, f1=0.6055045871559633,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023052005097270012
step: 10, loss: 0.00130128872115165
step: 20, loss: 0.0002233231207355857
step: 30, loss: 0.0001462176878703758
step: 40, loss: 0.07762588560581207
step: 50, loss: 0.00045536100515164435
step: 60, loss: 0.024047518149018288
step: 70, loss: 0.0007376528228633106
step: 80, loss: 0.0008078819955699146
step: 90, loss: 0.006059459410607815
step: 100, loss: 0.0015412616776302457
step: 110, loss: 0.0012013037921860814
step: 120, loss: 0.00019617094949353486
step: 130, loss: 0.002009339863434434
step: 140, loss: 0.0001594608329469338
step: 150, loss: 0.0007085895631462336
step: 160, loss: 0.0006734480848535895
step: 170, loss: 0.00042602611938491464
step: 180, loss: 0.0025560841895639896
step: 190, loss: 0.00017581565771251917
step: 200, loss: 0.00018492071831133217
step: 210, loss: 0.011802962981164455
/usr/local/lib/python3.7/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=0.5610859728506787, f1=0.6000000000000001,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006277641747146845
step: 10, loss: 0.00011492678459035233
step: 20, loss: 0.022745702415704727
step: 30, loss: 0.020457861945033073
step: 40, loss: 0.0013349531218409538
step: 50, loss: 0.00047090614680200815
step: 60, loss: 0.00042197867878712714
step: 70, loss: 0.011088176630437374
step: 80, loss: 0.0005682246992364526

```

```

step: 90, loss: 0.0001010531050269492
step: 100, loss: 0.019743507727980614
step: 110, loss: 0.00017991459753829986
step: 120, loss: 0.00027770333690568805
step: 130, loss: 0.00021787370496895164
step: 140, loss: 0.007059624418616295
step: 150, loss: 0.0026060377713292837
step: 160, loss: 0.029601560905575752
step: 170, loss: 9.628834231989458e-05
step: 180, loss: 0.0005278571625240147
step: 190, loss: 0.0002802237286232412
step: 200, loss: 0.029836630448698997
step: 210, loss: 0.0005498878890648484
/usr/local/lib/python3.7/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=0.5583524027459954, f1=0.5995423340961099,
best_f1=0.6452905811623246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005759702762588859
step: 10, loss: 0.016410116106271744
step: 20, loss: 0.0001626913290238008
step: 30, loss: 0.00023825163953006268
step: 40, loss: 0.00013031532580498606
step: 50, loss: 0.01171413529664278
step: 60, loss: 0.00029741256730630994
step: 70, loss: 0.0002730275155045092
step: 80, loss: 0.00013549065624829382
step: 90, loss: 0.0007366005447693169
step: 100, loss: 0.00036201823968440294
step: 110, loss: 0.00029428728157654405
step: 120, loss: 0.0011050974717363715
step: 130, loss: 0.00016153468459378928
step: 140, loss: 0.039376381784677505
step: 150, loss: 0.005427696742117405
step: 160, loss: 0.007764906622469425
step: 170, loss: 0.0002736827591434121
step: 180, loss: 0.0002998471027240157
step: 190, loss: 0.00043299520621076226
step: 200, loss: 0.0003677497443277389
step: 210, loss: 0.0003980435139965266
/usr/local/lib/python3.7/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=0.5492957746478874, f1=0.5813953488372092,
best_f1=0.6452905811623246

```

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

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

```
step: 0, loss: 0.0002164265315514058
step: 10, loss: 0.0001884384109871462
step: 20, loss: 0.0005536008975468576
step: 30, loss: 0.01617651991546154
step: 40, loss: 0.009819394908845425
step: 50, loss: 9.852898074313998e-05
step: 60, loss: 0.018809763714671135
step: 70, loss: 0.0026279401499778032
step: 80, loss: 0.00017606645997148007
step: 90, loss: 0.00046493211993947625
step: 100, loss: 0.020348727703094482
step: 110, loss: 0.00011077075760113075
step: 120, loss: 0.00011666442878777161
step: 130, loss: 0.00032041294616647065
step: 140, loss: 0.00010679856495698914
step: 150, loss: 0.0003619845083449036
step: 160, loss: 0.0008952996577136219
step: 170, loss: 0.0025686249136924744
step: 180, loss: 0.00017520504479762167
step: 190, loss: 0.00010419584577903152
step: 200, loss: 0.0008669685339555144
step: 210, loss: 0.0004135944473091513
```

```
/usr/local/lib/python3.7/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=0.5560747663551402, f1=0.586046511627907,
```

```
best_f1=0.6452905811623246
```

```
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 [00:08, 281.27it/s]
load_f1 = 0.6202783300198806
real_f1 = 0.616
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.71it/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 bert-base-uncased \
    --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 bert-base-uncased \

```

```
--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: 0.5702155828475952  
step: 10, loss: 0.3836224675178528  
step: 20, loss: 0.29920610785484314  
step: 30, loss: 0.416873961687088  
step: 40, loss: 0.4372500479221344  
step: 50, loss: 0.33031418919563293  
step: 60, loss: 0.274509996175766
```

```

step: 70, loss: 0.24865178763866425
step: 80, loss: 0.2500951290130615
step: 90, loss: 0.311379998922348
step: 100, loss: 0.2898648977279663
step: 110, loss: 0.3344440758228302
step: 120, loss: 0.09324097633361816
step: 130, loss: 0.08567391335964203
step: 140, loss: 0.03678925707936287
step: 150, loss: 0.10789704322814941
step: 160, loss: 0.14984479546546936
step: 170, loss: 0.20682837069034576
step: 180, loss: 0.01976926252245903
step: 190, loss: 0.1942034661769867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6826923076923076, f1=0.7153284671532846,
best_f1=0.7153284671532846
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2956336736679077
step: 10, loss: 0.031310100108385086
step: 20, loss: 0.12008827924728394
step: 30, loss: 0.1635238528251648
step: 40, loss: 0.13057096302509308
step: 50, loss: 0.043385252356529236
step: 60, loss: 0.24106577038764954
step: 70, loss: 0.3683586120605469
step: 80, loss: 0.168092280626297
step: 90, loss: 0.21834388375282288
step: 100, loss: 0.020674103870987892
step: 110, loss: 0.2171851247549057
step: 120, loss: 0.17850612103939056
step: 130, loss: 0.06762033700942993
step: 140, loss: 0.09340450912714005
step: 150, loss: 0.08743634819984436
step: 160, loss: 0.06408315151929855
step: 170, loss: 0.1396813839673996
step: 180, loss: 0.13979472219944
step: 190, loss: 0.0459727980196476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7923497267759563, f1=0.8022284122562673,
best_f1=0.8022284122562673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08057229965925217
step: 10, loss: 0.16307112574577332
step: 20, loss: 0.07401979714632034
step: 30, loss: 0.09704585373401642
step: 40, loss: 0.03035420924425125
step: 50, loss: 0.12583871185779572
step: 60, loss: 0.024735018610954285
step: 70, loss: 0.08353149890899658
step: 80, loss: 0.14219403266906738
step: 90, loss: 0.151479110121727
step: 100, loss: 0.030592795461416245
step: 110, loss: 0.09082714468240738
step: 120, loss: 0.07148759067058563
step: 130, loss: 0.01106941606849432
step: 140, loss: 0.04693306237459183
step: 150, loss: 0.10662461072206497
step: 160, loss: 0.015533359721302986
step: 170, loss: 0.10966257750988007
step: 180, loss: 0.029515083879232407
step: 190, loss: 0.12080565840005875
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7958115183246073, f1=0.7851458885941645,
best_f1=0.7851458885941645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06044350191950798
step: 10, loss: 0.05021367594599724
step: 20, loss: 0.08799106627702713
step: 30, loss: 0.012075376696884632
step: 40, loss: 0.053986094892024994
step: 50, loss: 0.035611171275377274
step: 60, loss: 0.22459357976913452

```



```

step: 70, loss: 0.13074442744255066
step: 80, loss: 0.07893980294466019
step: 90, loss: 0.01019175536930561
step: 100, loss: 0.01481155026704073
step: 110, loss: 0.005755351856350899
step: 120, loss: 0.019682660698890686
step: 130, loss: 0.05625030770897865
step: 140, loss: 0.01275554858148098
step: 150, loss: 0.01053585298359394
step: 160, loss: 0.037732526659965515
step: 170, loss: 0.1378309279680252
step: 180, loss: 0.012998834252357483
step: 190, loss: 0.16423946619033813
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.7774647887323942,
best_f1=0.7774647887323942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08114514499902725
step: 10, loss: 0.041523829102516174
step: 20, loss: 0.0769176259636879
step: 30, loss: 0.007329241372644901
step: 40, loss: 0.07311122864484787
step: 50, loss: 0.20590655505657196
step: 60, loss: 0.1386532485485077
step: 70, loss: 0.0037273357156664133
step: 80, loss: 0.014444788917899132
step: 90, loss: 0.07222390919923782
step: 100, loss: 0.0014577771071344614
step: 110, loss: 0.005960146430879831
step: 120, loss: 0.005933031905442476
step: 130, loss: 0.02347630448639393
step: 140, loss: 0.010341061279177666
step: 150, loss: 0.023266294971108437
step: 160, loss: 0.0508766770362854
step: 170, loss: 0.008281505666673183
step: 180, loss: 0.08159155398607254
step: 190, loss: 0.1985924243927002
/usr/local/lib/python3.7/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=0.7842105263157894, f1=0.7532467532467532,
best_f1=0.7774647887323942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034222134854644537
step: 10, loss: 0.04718136042356491
step: 20, loss: 0.0130492327734828
step: 30, loss: 0.001444653607904911
step: 40, loss: 0.005945586133748293
step: 50, loss: 0.015236676670610905
step: 60, loss: 0.0013912888243794441
step: 70, loss: 0.008568077348172665
step: 80, loss: 0.07872714847326279
step: 90, loss: 0.08581099659204483
step: 100, loss: 0.002056194469332695
step: 110, loss: 0.0033763523679226637
step: 120, loss: 0.03185780718922615
step: 130, loss: 0.07693180441856384
step: 140, loss: 0.023033302277326584
step: 150, loss: 0.0023629572242498398
step: 160, loss: 0.004261190537363291
step: 170, loss: 0.09295755624771118
step: 180, loss: 0.006278639659285545
step: 190, loss: 0.014068116433918476
/usr/local/lib/python3.7/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=0.7833333333333333, f1=0.7999999999999999,
best_f1=0.7774647887323942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002032299293205142
step: 10, loss: 0.008091830648481846
step: 20, loss: 0.005830393638461828
step: 30, loss: 0.0023516362998634577
step: 40, loss: 0.0016847988590598106
step: 50, loss: 0.001602275762706995
step: 60, loss: 0.009783570654690266
step: 70, loss: 0.001672651618719101
step: 80, loss: 0.05229629576206207
step: 90, loss: 0.006804086733609438
step: 100, loss: 0.0008148199412971735
step: 110, loss: 0.001307718106545508
step: 120, loss: 0.0069845011457800865
step: 130, loss: 0.001879932009615004
step: 140, loss: 0.0043743629939854145

```

```

step: 150, loss: 0.0014881929382681847
step: 160, loss: 0.15961001813411713
step: 170, loss: 0.003599154995754361
step: 180, loss: 0.00361996004357934
step: 190, loss: 0.0012383178109303117
/usr/local/lib/python3.7/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=0.7946666666666666, f1=0.7956989247311829,
best_f1=0.7774647887323942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010801676660776138
step: 10, loss: 0.0027001709677278996
step: 20, loss: 0.0011746302479878068
step: 30, loss: 0.0024364925920963287
step: 40, loss: 0.0009644146775826812
step: 50, loss: 0.0008555327076464891
step: 60, loss: 0.00562919769436121
step: 70, loss: 0.016477515920996666
step: 80, loss: 0.00019424264610279351
step: 90, loss: 0.015790782868862152
step: 100, loss: 0.08704125136137009
step: 110, loss: 0.0012244138633832335
step: 120, loss: 0.00041784142376855016
step: 130, loss: 0.0005228723748587072
step: 140, loss: 0.0024311027955263853
step: 150, loss: 0.0008802110678516328
step: 160, loss: 0.0007375647546723485
step: 170, loss: 0.0019694501534104347
step: 180, loss: 0.00491564953699708
step: 190, loss: 0.00269674276933074
/usr/local/lib/python3.7/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=0.7795698924731183, f1=0.7765957446808511,
best_f1=0.7774647887323942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003129478543996811
step: 10, loss: 0.028208276256918907
step: 20, loss: 0.001205503474920988
step: 30, loss: 0.00232895789667964
step: 40, loss: 0.000269347830908373
step: 50, loss: 0.009156467393040657
step: 60, loss: 0.004984751343727112

```

```

step: 70, loss: 0.0004167104489170015
step: 80, loss: 0.0005010993336327374
step: 90, loss: 0.0019458180759102106
step: 100, loss: 0.0013845915673300624
step: 110, loss: 0.0005960778798907995
step: 120, loss: 0.007905753329396248
step: 130, loss: 0.002773771295323968
step: 140, loss: 0.18314428627490997
step: 150, loss: 0.0026812897995114326
step: 160, loss: 0.006079643499106169
step: 170, loss: 0.016497069969773293
step: 180, loss: 0.0010350544471293688
step: 190, loss: 0.0011524721048772335
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8041775456919059, f1=0.7917737789203083,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033793009351938963
step: 10, loss: 0.0033078899141401052
step: 20, loss: 0.0013476446038112044
step: 30, loss: 0.0002863405679818243
step: 40, loss: 0.018343951553106308
step: 50, loss: 0.0007257536053657532
step: 60, loss: 0.0003480835584923625
step: 70, loss: 0.0004372376424726099
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.0005128409247845411
step: 90, loss: 0.001565228565596044
step: 100, loss: 0.0007517257472500205
step: 110, loss: 0.008999078534543514
step: 120, loss: 0.12250525504350662
step: 130, loss: 0.0002805425028782338
step: 140, loss: 0.0033096643164753914
step: 150, loss: 0.0011196006089448929
step: 160, loss: 0.0005305794766172767
step: 170, loss: 0.00022910661937203258
step: 180, loss: 0.000128921601572074
step: 190, loss: 0.0003266753046773374
/usr/local/lib/python3.7/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=0.7894736842105263, f1=0.7947368421052631,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011483923299238086
step: 10, loss: 0.00031188817229121923
step: 20, loss: 0.0017020761733874679
step: 30, loss: 0.0013549355790019035
step: 40, loss: 0.00012311217142269015
step: 50, loss: 0.0003417113039176911
step: 60, loss: 0.00037323770811781287
step: 70, loss: 0.0007866558735258877
step: 80, loss: 0.0005814250325784087
step: 90, loss: 0.00010504276724532247
step: 100, loss: 0.00017000663501676172
step: 110, loss: 0.00015951784735079855
step: 120, loss: 0.0005004038102924824
step: 130, loss: 0.0002187734644394368
step: 140, loss: 0.0004936994519084692
step: 150, loss: 0.00014902283146511763
step: 160, loss: 0.0006982521736063063
step: 170, loss: 0.0003652212326414883
step: 180, loss: 0.00022904294019099325
step: 190, loss: 0.0021641463972628117
/usr/local/lib/python3.7/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=0.7868852459016394, f1=0.8130081300813009,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014890571765135974
step: 10, loss: 0.005604640115052462
step: 20, loss: 0.00013102895172778517
step: 30, loss: 0.0004554119077511132
step: 40, loss: 0.00017773172294255346
step: 50, loss: 0.005759811028838158
step: 60, loss: 0.00032786044175736606
step: 70, loss: 9.898316784529015e-05
step: 80, loss: 0.00010955319157801569
step: 90, loss: 0.00027169616078026593
step: 100, loss: 0.0001742802414810285
step: 110, loss: 0.00018857847317121923
step: 120, loss: 0.0005595466936938465
step: 130, loss: 0.0003359736583661288

```

```

step: 140, loss: 0.0009906981140375137
step: 150, loss: 0.0004599730600602925
step: 160, loss: 0.0002258413878735155
step: 170, loss: 0.028764907270669937
step: 180, loss: 5.382017116062343e-05
step: 190, loss: 0.00010667372407624498
/usr/local/lib/python3.7/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=0.7878787878787877, f1=0.8199445983379501,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006206862162798643
step: 10, loss: 9.190778655465692e-05
step: 20, loss: 0.0005552738439291716
step: 30, loss: 0.0005100017879158258
step: 40, loss: 0.0004875875893048942
step: 50, loss: 0.0009482991881668568
step: 60, loss: 8.689275273354724e-05
step: 70, loss: 0.004376382566988468
step: 80, loss: 0.05556127056479454
step: 90, loss: 0.0004548032011371106
step: 100, loss: 0.019768504425883293
step: 110, loss: 0.000327251706039533
step: 120, loss: 0.002163359895348549
step: 130, loss: 8.758237527217716e-05
step: 140, loss: 0.0001855078589869663
step: 150, loss: 0.00030200841138139367
step: 160, loss: 0.0002241660258732736
step: 170, loss: 0.0006248675053939223
step: 180, loss: 0.00012566331133712083
step: 190, loss: 0.08425241708755493
/usr/local/lib/python3.7/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=0.7845303867403315, f1=0.8044077134986225,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013595340715255588
step: 10, loss: 0.00017399652278982103
step: 20, loss: 0.00019687198800966144
step: 30, loss: 0.00011040570097975433
step: 40, loss: 0.0002890509204007685
step: 50, loss: 0.00045038352254778147

```

```

step: 60, loss: 0.0003403876326046884
step: 70, loss: 0.00020084828429389745
step: 80, loss: 0.00011068554886151105
step: 90, loss: 0.00035833672154694796
step: 100, loss: 0.0018613828578963876
step: 110, loss: 0.00016717828111723065
step: 120, loss: 0.00023862652597017586
step: 130, loss: 0.0001663265866227448
step: 140, loss: 0.00014619404100812972
step: 150, loss: 0.00016459856124129146
step: 160, loss: 0.0005644951015710831
step: 170, loss: 0.00019694012007676065
step: 180, loss: 0.00020923679403495044
step: 190, loss: 0.0011075170477852225
/usr/local/lib/python3.7/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=0.7927461139896373, f1=0.7948051948051948,
best_f1=0.7917737789203083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006905566435307264
step: 10, loss: 8.728742250241339e-05
step: 20, loss: 0.00016162953397724777
step: 30, loss: 0.00011198624997632578
step: 40, loss: 0.00039611064130440354
step: 50, loss: 0.00047178970999084413
step: 60, loss: 0.00019262643763795495
step: 70, loss: 0.0012662302469834685
step: 80, loss: 0.0016822739271447062
step: 90, loss: 0.00010025635856436566
step: 100, loss: 0.002405765699222684
step: 110, loss: 0.00010481294157216325
step: 120, loss: 0.00021389633184298873
step: 130, loss: 0.0002198223810410127
step: 140, loss: 0.0008918584208004177
step: 150, loss: 0.00015531318786088377
step: 160, loss: 0.00014494686911348253
step: 170, loss: 0.00026655549299903214
step: 180, loss: 0.0004924798267893493
step: 190, loss: 0.0002608966315165162
/usr/local/lib/python3.7/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=0.7916666666666666, f1=0.7905759162303665,
best_f1=0.7917737789203083
mv: cannot stat '*_dev.pt': No such file or directory

```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
2049it [00:12, 167.51it/s]
load_f1 = 0.8053333333333333
real_f1 = 0.7989417989417988
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, 183.53it/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 bert-base-uncased \
--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 bert-base-uncased \
  --use_gpu \
  --fp16 \
  --checkpoint_path checkpoints/

```

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

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
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: 0.6196228861808777
step: 10, loss: 0.39933598041534424
step: 20, loss: 0.30724263191223145
step: 30, loss: 0.38219794631004333
step: 40, loss: 0.2735332250595093
step: 50, loss: 0.2742752134799957
step: 60, loss: 0.2130870223045349
step: 70, loss: 0.3485220968723297
step: 80, loss: 0.29464730620384216
step: 90, loss: 0.2597172260284424
step: 100, loss: 0.21661250293254852
step: 110, loss: 0.2502395510673523
step: 120, loss: 0.17414219677448273
step: 130, loss: 0.022846849635243416
step: 140, loss: 0.18572096526622772
step: 150, loss: 0.28571823239326477
step: 160, loss: 0.09379509091377258
step: 170, loss: 0.2237543910741806
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7276887871853546, f1=0.7136363636363636,
best_f1=0.7136363636363636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06043141707777977
step: 10, loss: 0.3401145339012146
step: 20, loss: 0.1467931866645813
step: 30, loss: 0.17829269170761108
step: 40, loss: 0.13352158665657043
step: 50, loss: 0.10286816209554672
```

```

step: 60, loss: 0.07268332690000534
step: 70, loss: 0.084718719124794
step: 80, loss: 0.13715822994709015
step: 90, loss: 0.14804226160049438
step: 100, loss: 0.07689683884382248
step: 110, loss: 0.04126761108636856
step: 120, loss: 0.056172922253608704
step: 130, loss: 0.12318399548530579
step: 140, loss: 0.36141064763069153
step: 150, loss: 0.17943806946277618
step: 160, loss: 0.1454332023859024
step: 170, loss: 0.04997464269399643
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7461139896373058, f1=0.7860696517412936,
best_f1=0.7860696517412936
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07172627002000809
step: 10, loss: 0.05240214616060257
step: 20, loss: 0.08119945973157883
step: 30, loss: 0.15223455429077148
step: 40, loss: 0.17177042365074158
step: 50, loss: 0.1756497323513031
step: 60, loss: 0.03955001384019852
step: 70, loss: 0.0593859925866127
step: 80, loss: 0.07488259673118591
step: 90, loss: 0.16890372335910797
step: 100, loss: 0.0034857902210205793
step: 110, loss: 0.13568605482578278
step: 120, loss: 0.05946598947048187
step: 130, loss: 0.1291753202676773
step: 140, loss: 0.05205842852592468
step: 150, loss: 0.010964316315948963
step: 160, loss: 0.012530962936580181
step: 170, loss: 0.032856933772563934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7669902912621359, f1=0.7848699763593381,
best_f1=0.7848699763593381
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01890941895544529
step: 10, loss: 0.03478655591607094
step: 20, loss: 0.049054067581892014
step: 30, loss: 0.020390208810567856
step: 40, loss: 0.0011237503495067358
step: 50, loss: 0.08822202682495117
step: 60, loss: 0.1810912787914276
step: 70, loss: 0.003396302228793502
step: 80, loss: 0.05588885769248009
step: 90, loss: 0.03993607684969902
step: 100, loss: 0.05111761391162872
step: 110, loss: 0.03405589237809181
step: 120, loss: 0.01167586725205183
step: 130, loss: 0.0208131093531847
step: 140, loss: 0.11245584487915039
step: 150, loss: 0.22941400110721588
step: 160, loss: 0.06553412973880768
step: 170, loss: 0.0037049211096018553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7917737789203084, f1=0.7951219512195121,
best_f1=0.7951219512195121
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.019348373636603355
step: 10, loss: 0.0047509875148534775
step: 20, loss: 0.06311140954494476
step: 30, loss: 0.05799786373972893
step: 40, loss: 0.006668620742857456
step: 50, loss: 0.08698318898677826
step: 60, loss: 0.026570696383714676
step: 70, loss: 0.07476731389760971
step: 80, loss: 0.04302487149834633
step: 90, loss: 0.1427266150712967
step: 100, loss: 0.03599507734179497
step: 110, loss: 0.00778178870677948

```

```

step: 120, loss: 0.045794423669576645
step: 130, loss: 0.031062986701726913
step: 140, loss: 0.0029855507891625166
step: 150, loss: 0.043681707233190536
step: 160, loss: 0.009345878846943378
step: 170, loss: 0.012876022607088089
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8163265306122449, f1=0.801007556675063,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024349349550902843
step: 10, loss: 0.01246707420796156
step: 20, loss: 0.008436968550086021
step: 30, loss: 0.00366060808300972
step: 40, loss: 0.05145813897252083
step: 50, loss: 0.33885252475738525
step: 60, loss: 0.014850914478302002
step: 70, loss: 0.10497765243053436
step: 80, loss: 0.019096823409199715
step: 90, loss: 0.03112824447453022
step: 100, loss: 0.016238398849964142
step: 110, loss: 0.0008717470336705446
step: 120, loss: 0.003292483976110816
step: 130, loss: 0.0006960314349271357
step: 140, loss: 0.0009502628818154335
step: 150, loss: 0.13097599148750305
step: 160, loss: 0.08523133397102356
step: 170, loss: 0.009932897053658962
/usr/local/lib/python3.7/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=0.7979539641943735, f1=0.8040201005025125,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019195096101611853
step: 10, loss: 0.013225359842181206
step: 20, loss: 0.019429627805948257
step: 30, loss: 0.02416081354022026

```

```

step: 40, loss: 0.0013128925347700715
step: 50, loss: 0.022173311561346054
step: 60, loss: 0.00229705311357975
step: 70, loss: 0.0003211093135178089
step: 80, loss: 0.00779046630486846
step: 90, loss: 0.0002247638622066006
step: 100, loss: 0.0009710636222735047
step: 110, loss: 0.004248865880072117
step: 120, loss: 0.005733162630349398
step: 130, loss: 0.10137295722961426
step: 140, loss: 0.0012844894081354141
step: 150, loss: 0.0025098284240812063
step: 160, loss: 0.0007781382300890982
step: 170, loss: 0.032245833426713943
/usr/local/lib/python3.7/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=0.7882037533512065, f1=0.8148148148148149,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01392060425132513
step: 10, loss: 0.0003236115735489875
step: 20, loss: 0.12130532413721085
step: 30, loss: 0.011273130774497986
step: 40, loss: 0.00019934013835154474
step: 50, loss: 0.03867615759372711
step: 60, loss: 0.0005625950288958848
step: 70, loss: 0.00039867445593699813
step: 80, loss: 0.0030711544677615166
step: 90, loss: 0.002966040512546897
step: 100, loss: 0.014893629588186741
step: 110, loss: 0.10648414492607117
step: 120, loss: 0.000481956172734499
step: 130, loss: 0.04728879779577255
step: 140, loss: 0.010034782811999321
step: 150, loss: 0.058347709476947784
step: 160, loss: 0.10581134259700775
step: 170, loss: 0.00539632560685277
/usr/local/lib/python3.7/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=0.7929292929292929, f1=0.7821782178217821,
best_f1=0.801007556675063
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0005739156622439623
step: 10, loss: 0.0008634804398752749
step: 20, loss: 0.00031753050279803574
step: 30, loss: 0.01611943170428276
step: 40, loss: 0.013466579839587212
step: 50, loss: 0.00023319771571550518
step: 60, loss: 0.0007529794238507748
step: 70, loss: 0.032060205936431885
step: 80, loss: 0.001007425133138895
step: 90, loss: 0.03324209526181221
step: 100, loss: 0.004220642615109682
step: 110, loss: 0.0004282856243662536
step: 120, loss: 0.03254341706633568
step: 130, loss: 0.042230598628520966
step: 140, loss: 0.00020105308794882149
step: 150, loss: 0.005991777870804071
step: 160, loss: 0.021100565791130066
step: 170, loss: 0.02218502201139927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8217054263565892, f1=0.7917737789203084,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022024255245923996
step: 10, loss: 0.00018385794828645885
step: 20, loss: 0.14623376727104187
step: 30, loss: 0.0063887485302984715
step: 40, loss: 0.00018589307728689164
step: 50, loss: 0.06394912302494049
step: 60, loss: 0.00015468805213458836
step: 70, loss: 0.004066022578626871
step: 80, loss: 0.0014900639653205872
step: 90, loss: 0.004190411418676376
step: 100, loss: 0.00043010921217501163
step: 110, loss: 0.00033321548835374415
step: 120, loss: 0.00024392161867581308
step: 130, loss: 0.010018400847911835
step: 140, loss: 0.004200671333819628
step: 150, loss: 0.0038411584682762623
step: 160, loss: 0.010998778976500034
step: 170, loss: 0.00015542608161922544

```

```

/usr/local/lib/python3.7/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=0.7897435897435897,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04774194583296776
step: 10, loss: 0.005075046326965094
step: 20, loss: 0.00019204682030249387
step: 30, loss: 0.00010329913493478671
step: 40, loss: 0.00041238105040974915
step: 50, loss: 0.006517794914543629
step: 60, loss: 0.01490430161356926
step: 70, loss: 0.00011733171413652599
step: 80, loss: 0.0001263357262359932
step: 90, loss: 0.00014533119974657893
step: 100, loss: 0.0001305321929976344
step: 110, loss: 0.013237877748906612
step: 120, loss: 0.0005859804805368185
step: 130, loss: 0.00013198982924222946
step: 140, loss: 0.00012130050163250417
step: 150, loss: 0.056329019367694855
step: 160, loss: 0.010028978809714317
step: 170, loss: 0.0004444525111466646
/usr/local/lib/python3.7/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=0.7912621359223301, f1=0.7761904761904763,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014674892649054527
step: 10, loss: 0.00038714526453986764
step: 20, loss: 0.00010915351595031098
step: 30, loss: 0.00011508540774229914
step: 40, loss: 0.00026142876595258713
step: 50, loss: 0.00011893634655280039
step: 60, loss: 0.0005742667708545923
step: 70, loss: 0.06378734111785889
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 80, loss: 9.830964700086042e-05
step: 90, loss: 0.00013999029761180282
step: 100, loss: 0.00023496380890719593
step: 110, loss: 0.0002464156423229724
step: 120, loss: 0.0002303333458257839

```



```

step: 130, loss: 0.02107195369899273
step: 140, loss: 0.029963472858071327
step: 150, loss: 0.0037411802913993597
step: 160, loss: 0.00014745022053830326
step: 170, loss: 9.574176510795951e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.801007556675063, f1=0.7654320987654322,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022442620247602463
step: 10, loss: 0.0001006435850285925
step: 20, loss: 0.00017724369536153972
step: 30, loss: 6.731012399541214e-05
step: 40, loss: 0.00015562308544758707
step: 50, loss: 0.00010967948037432507
step: 60, loss: 0.0008317791507579386
step: 70, loss: 0.00023919122759252787
step: 80, loss: 0.0003892956010531634
step: 90, loss: 7.414602441713214e-05
step: 100, loss: 0.00023653706011828035
step: 110, loss: 7.376277062576264e-05
step: 120, loss: 0.03564397245645523
step: 130, loss: 0.00011880516831297427
step: 140, loss: 0.00039788498543202877
step: 150, loss: 0.0032705513294786215
step: 160, loss: 0.0008479305543005466
step: 170, loss: 0.00042022368870675564
/usr/local/lib/python3.7/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=0.8064516129032259, f1=0.8020833333333334,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014902418479323387
step: 10, loss: 7.680935959797353e-05
step: 20, loss: 0.00027412292547523975
step: 30, loss: 0.0002554076199885458
step: 40, loss: 9.886805491987616e-05
step: 50, loss: 0.00015462710871361196
step: 60, loss: 7.534863834735006e-05
step: 70, loss: 0.0006323034758679569
step: 80, loss: 0.0017808134434744716

```

```

step: 90, loss: 7.289768109330907e-05
step: 100, loss: 0.00027334579499438405
step: 110, loss: 0.00013981306983623654
step: 120, loss: 0.02641107887029648
step: 130, loss: 0.052360907196998596
step: 140, loss: 0.00010293009836459532
step: 150, loss: 0.0001853188732638955
step: 160, loss: 4.654255462810397e-05
step: 170, loss: 6.961715553188697e-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=0.8071979434447302, f1=0.7938931297709925,
best_f1=0.7917737789203084
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010869545076275244
step: 10, loss: 0.0003566346422303468
step: 20, loss: 0.00019642870756797493
step: 30, loss: 8.898483065422624e-05
step: 40, loss: 0.00015089927182998508
step: 50, loss: 6.512270920211449e-05
step: 60, loss: 0.00018356842338107526
step: 70, loss: 6.777067028451711e-05
step: 80, loss: 0.004825887270271778
step: 90, loss: 0.0017568438779562712
step: 100, loss: 0.0007957497728057206
step: 110, loss: 9.62048870860599e-05
step: 120, loss: 0.0005524765001609921
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 130, loss: 0.00012035373947583139
step: 140, loss: 0.00010769845539471135
step: 150, loss: 0.000209121557418257
step: 160, loss: 7.086442928994074e-05
step: 170, loss: 0.00011826871923403814
/usr/local/lib/python3.7/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=0.7946666666666666, f1=0.8020833333333333,
best_f1=0.7917737789203084
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

```

Defaults for this optimization level are:
enabled          : True
opt_level        : 02

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled              : True
opt_level            : 02
cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:08, 236.69it/s]
load_f1 = 0.8207792207792208
real_f1 = 0.8155844155844156
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:22, 194.83it/s]

```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

```

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

```

```

--lm bert-base-uncased \
--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 bert-base-uncased \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/

```

Downloading: 100% 433/433 [00:00<00:00, 649kB/s]
 Downloading: 100% 232k/232k [00:00<00:00, 4.07MB/s]
 Downloading: 100% 440M/440M [00:07<00:00, 56.9MB/s]
 Selected optimization level O2: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

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

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

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:136:
UserWarning: Detected call of `lr_scheduler.step()` before `optimizer.step()`.
In PyTorch 1.1.0 and later, you should call them in the opposite order:
`optimizer.step()` before `lr_scheduler.step()`. Failure to do this will result
in PyTorch skipping the first value of the learning rate schedule. See more
details at https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-
rate
  "https://pytorch.org/docs/stable/optim.html#how-to-adjust-learning-rate",
UserWarning)
step: 0, loss: 0.6203136444091797
step: 10, loss: 0.6090444922447205
step: 20, loss: 0.49729305505752563
step: 30, loss: 0.26589420437812805
step: 40, loss: 0.2327870875597
step: 50, loss: 0.07982302457094193
step: 60, loss: 0.10154690593481064
step: 70, loss: 0.10197766125202179
step: 80, loss: 0.10242507606744766
step: 90, loss: 0.1576164960861206
step: 100, loss: 0.023660091683268547
step: 110, loss: 0.11278767138719559
step: 120, loss: 0.00947610568255186
step: 130, loss: 0.012313169427216053
step: 140, loss: 0.008549846708774567
step: 150, loss: 0.027086645364761353
step: 160, loss: 0.005439393687993288
step: 170, loss: 0.18384641408920288
step: 180, loss: 0.012912136502563953
step: 190, loss: 0.10595869272947311
step: 200, loss: 0.10133808851242065
step: 210, loss: 0.007792278658598661
step: 220, loss: 0.038231950253248215
step: 230, loss: 0.025000110268592834
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9663677130044843, f1=0.9695603156708005,
best_f1=0.9695603156708005
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006685005035251379
step: 10, loss: 0.06122604385018349
step: 20, loss: 0.04237993806600571

```

```

step: 30, loss: 0.20470644533634186
step: 40, loss: 0.07428956776857376
step: 50, loss: 0.007970769889652729
step: 60, loss: 0.006369869224727154
step: 70, loss: 0.06876637041568756
step: 80, loss: 0.017426123842597008
step: 90, loss: 0.0028628273867070675
step: 100, loss: 0.1038855636119843
step: 110, loss: 0.16532394289970398
step: 120, loss: 0.12045326828956604
step: 130, loss: 0.0095136109739542
step: 140, loss: 0.005261925049126148
step: 150, loss: 0.025200532749295235
step: 160, loss: 0.014332551509141922
step: 170, loss: 0.0029257687274366617
step: 180, loss: 0.031022600829601288
step: 190, loss: 0.002776074456050992
step: 200, loss: 0.007127845194190741
step: 210, loss: 0.007954521104693413
step: 220, loss: 0.06550294160842896
step: 230, loss: 0.0045699626207351685
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not 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 2: dev_f1=0.9732739420935412, f1=0.9696969696969697,
best_f1=0.9696969696969697
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01039542444050312
step: 10, loss: 0.005203761626034975
step: 20, loss: 0.0027406783774495125
step: 30, loss: 0.001264840946532786
step: 40, loss: 0.13681627810001373
step: 50, loss: 0.0025718004908412695
step: 60, loss: 0.005658379755914211
step: 70, loss: 0.09106779098510742
step: 80, loss: 0.0011980487033724785
step: 90, loss: 0.003239715239033103
step: 100, loss: 0.001729230978526175
step: 110, loss: 0.0016505997627973557
step: 120, loss: 0.00333068217150867
step: 130, loss: 0.0023095218930393457
step: 140, loss: 0.0011402072850614786

```

```

step: 150, loss: 0.010353309102356434
step: 160, loss: 0.052606381475925446
step: 170, loss: 0.012831867672502995
step: 180, loss: 0.052789632230997086
step: 190, loss: 0.0054565249010920525
step: 200, loss: 0.005081301089376211
step: 210, loss: 0.00681639788672328
step: 220, loss: 0.0013100765645503998
step: 230, loss: 0.0009961072355508804
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.975609756097561, f1=0.9742441209406495,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007054973975755274
step: 10, loss: 0.0012748984154313803
step: 20, loss: 0.0017985348822548985
step: 30, loss: 0.0008376381592825055
step: 40, loss: 0.0016268648905679584
step: 50, loss: 0.0003530631074681878
step: 60, loss: 0.0015560067258775234
step: 70, loss: 0.0014265121426433325
step: 80, loss: 0.004743081517517567
step: 90, loss: 0.0017458280781283975
step: 100, loss: 0.00036667747190222144
step: 110, loss: 0.00036624615313485265
step: 120, loss: 0.028861388564109802
step: 130, loss: 0.000830171920824796
step: 140, loss: 0.004536097403615713
step: 150, loss: 0.06854332238435745
step: 160, loss: 0.03915450721979141
step: 170, loss: 0.006158101838082075
step: 180, loss: 0.0004778072179760784
step: 190, loss: 0.0020205934997648
step: 200, loss: 0.004174020141363144
step: 210, loss: 0.14743784070014954
step: 220, loss: 0.0007484497618861496
step: 230, loss: 0.006114515941590071
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.968609865470852,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011089156614616513
step: 10, loss: 0.030930452048778534
step: 20, loss: 0.0013225848088040948
step: 30, loss: 0.0009122778428718448
step: 40, loss: 0.0008043685229495168
step: 50, loss: 0.04201795160770416
step: 60, loss: 0.0017717125592753291
step: 70, loss: 0.0006277997163124382
step: 80, loss: 0.0029925189446657896
step: 90, loss: 0.0025745525490492582
step: 100, loss: 0.00028829669463448226
step: 110, loss: 0.0005890181055292487
step: 120, loss: 0.0002635337586980313
step: 130, loss: 0.11814319342374802
step: 140, loss: 0.025061678141355515
step: 150, loss: 0.0024781005922704935
step: 160, loss: 0.012722463347017765
step: 170, loss: 0.0026803649961948395
step: 180, loss: 0.00022198984515853226
step: 190, loss: 0.04063091799616814
step: 200, loss: 0.0008623615140095353
step: 210, loss: 0.00038147493614815176
step: 220, loss: 0.09630849957466125
step: 230, loss: 0.00050015514716506
/usr/local/lib/python3.7/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=0.9745293466223698, f1=0.9776286353467561,
best_f1=0.9742441209406495
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010858844965696335
step: 10, loss: 0.005254772491753101
step: 20, loss: 0.013450803235173225
step: 30, loss: 0.00027963309548795223
step: 40, loss: 0.00047587064909748733
step: 50, loss: 0.0007280459976755083
step: 60, loss: 0.00021749631559941918
step: 70, loss: 0.0154136186465621
step: 80, loss: 0.0007688934565521777
step: 90, loss: 0.0004805296193808317
step: 100, loss: 0.01650143787264824

```



```

step: 110, loss: 0.0011751118581742048
step: 120, loss: 0.00311021111097425222
step: 130, loss: 0.000861030537635088
step: 140, loss: 0.00040979532059282064
step: 150, loss: 0.005370547994971275
step: 160, loss: 0.0010284302989020944
step: 170, loss: 0.0004058251215610653
step: 180, loss: 0.07008382678031921
step: 190, loss: 0.002156592905521393
step: 200, loss: 0.00018153547716792673
step: 210, loss: 0.00025688466848805547
step: 220, loss: 0.0001417480962118134
step: 230, loss: 0.029992982745170593
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9766925638179801, f1=0.9764837625979844,
best_f1=0.9764837625979844
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03792852535843849
step: 10, loss: 0.00013172619219403714
step: 20, loss: 0.00022896949667483568
step: 30, loss: 0.00016714679077267647
step: 40, loss: 0.001472764415666461
step: 50, loss: 0.0003106906951870769
step: 60, loss: 0.007000145502388477
step: 70, loss: 0.0024141264148056507
step: 80, loss: 0.0036908115725964308
step: 90, loss: 0.001114372629672289
step: 100, loss: 0.00044169844477437437
step: 110, loss: 0.00022299677948467433
step: 120, loss: 0.0009932825341820717
step: 130, loss: 0.00046735629439353943
step: 140, loss: 0.006286637857556343
step: 150, loss: 0.0015910181682556868
step: 160, loss: 0.11693967133760452
step: 170, loss: 0.007845073938369751
step: 180, loss: 0.0014428726863116026
step: 190, loss: 0.0004969812580384314
step: 200, loss: 0.0519932359457016
step: 210, loss: 0.00025646568974480033
step: 220, loss: 0.0007134417537599802

```

```

step: 230, loss: 0.0008424681727774441
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9788182831661093, f1=0.9755011135857461,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011436459608376026
step: 10, loss: 0.0018933292012661695
step: 20, loss: 0.0009468463831581175
step: 30, loss: 0.0001476348115829751
step: 40, loss: 0.0018967631040140986
step: 50, loss: 0.0013764846371486783
step: 60, loss: 0.00017890507297124714
step: 70, loss: 0.0002432514593238011
step: 80, loss: 0.0015375993680208921
step: 90, loss: 9.572722774464637e-05
step: 100, loss: 0.00026531759067438543
step: 110, loss: 0.0035335770808160305
step: 120, loss: 0.0005013828631490469
step: 130, loss: 0.001065199845470488
step: 140, loss: 0.0001143014378612861
step: 150, loss: 9.473761019762605e-05
step: 160, loss: 0.00013876651064492762
step: 170, loss: 8.66696864250116e-05
step: 180, loss: 0.00015495880506932735
step: 190, loss: 0.00020435960323084146
step: 200, loss: 0.002072185045108199
step: 210, loss: 0.0006966587388888001
step: 220, loss: 0.0011385338148102164
step: 230, loss: 0.0021629626862704754
/usr/local/lib/python3.7/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=0.9732142857142857, f1=0.9776785714285714,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007118204375728965
step: 10, loss: 0.00014521428965963423
step: 20, loss: 0.0005988253396935761

```

```

step: 30, loss: 0.00019733711087610573
step: 40, loss: 0.008111165836453438
step: 50, loss: 0.00013735094398725778
step: 60, loss: 0.00016871883417479694
step: 70, loss: 0.0003814975207205862
step: 80, loss: 0.00011552208889042959
step: 90, loss: 0.06812690198421478
step: 100, loss: 0.00040233368054032326
step: 110, loss: 8.909563621273264e-05
step: 120, loss: 0.00010414614371256903
step: 130, loss: 0.00016074338054750115
step: 140, loss: 0.00013115581532474607
step: 150, loss: 0.0003232079034205526
step: 160, loss: 0.0001597989903530106
step: 170, loss: 0.0003171678108628839
step: 180, loss: 0.004128125496208668
step: 190, loss: 5.4761905630584806e-05
step: 200, loss: 0.0007729026256129146
step: 210, loss: 0.00011584136518649757
step: 220, loss: 4.944383181282319e-05
step: 230, loss: 0.00015114012057892978
/usr/local/lib/python3.7/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=0.9751693002257337, f1=0.9785310734463276,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011174352403031662
step: 10, loss: 9.101275645662099e-05
step: 20, loss: 6.252714956644922e-05
step: 30, loss: 4.320436710258946e-05
step: 40, loss: 0.0005581480800174177
step: 50, loss: 6.20713472017087e-05
step: 60, loss: 4.084174724994227e-05
step: 70, loss: 0.00024602163466624916
step: 80, loss: 9.52604241319932e-05
step: 90, loss: 0.00011723016359610483
step: 100, loss: 0.00011928498861379921
step: 110, loss: 0.00014941337576601654
step: 120, loss: 0.0003884203324560076
step: 130, loss: 4.882452412857674e-05
step: 140, loss: 0.036131247878074646
step: 150, loss: 0.021745938807725906
step: 160, loss: 8.465252176392823e-05
step: 170, loss: 0.00018687530246097594
step: 180, loss: 0.00011865893611684442

```

```

step: 190, loss: 0.00011729911784641445
step: 200, loss: 3.245012339903042e-05
step: 210, loss: 7.958280184539035e-05
step: 220, loss: 0.00014964950969442725
step: 230, loss: 0.0001668405457166955
/usr/local/lib/python3.7/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=0.9719416386083053, f1=0.9832402234636871,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.822204497642815e-05
step: 10, loss: 0.0001755817502271384
step: 20, loss: 0.0003405751776881516
step: 30, loss: 0.02094799093902111
step: 40, loss: 0.001842131488956511
step: 50, loss: 0.001225697691552341
step: 60, loss: 0.0004038875922560692
step: 70, loss: 0.0008352554286830127
step: 80, loss: 5.6601933465572074e-05
step: 90, loss: 8.610571967437863e-05
step: 100, loss: 0.0001233528892043978
step: 110, loss: 0.00010024051152868196
step: 120, loss: 5.664722266374156e-05
step: 130, loss: 0.00012681575026363134
step: 140, loss: 0.0001598795352037996
step: 150, loss: 0.0003007844206877053
step: 160, loss: 5.1115879614371806e-05
step: 170, loss: 0.00014309374091681093
step: 180, loss: 8.768062252784148e-05
step: 190, loss: 0.00014711881522089243
step: 200, loss: 0.002121708123013377
step: 210, loss: 8.027219155337662e-05
step: 220, loss: 5.2252347813919187e-05
step: 230, loss: 6.773220957256854e-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=0.9753914988814317, f1=0.9775784753363228,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001236708485521376
step: 10, loss: 0.00012253034219611436
step: 20, loss: 0.0003580800839699805

```

```

step: 30, loss: 0.00021873302466701716
step: 40, loss: 0.00016614010382909328
step: 50, loss: 0.00041336423601023853
step: 60, loss: 0.00012046777555951849
step: 70, loss: 0.029461340978741646
step: 80, loss: 0.00020788866095244884
step: 90, loss: 8.683915075380355e-05
step: 100, loss: 4.3475880374899134e-05
step: 110, loss: 7.773165998514742e-05
step: 120, loss: 0.0011307810200378299
step: 130, loss: 3.9605940401088446e-05
step: 140, loss: 5.9409929235698655e-05
step: 150, loss: 6.316474173218012e-05
step: 160, loss: 9.545715874992311e-05
step: 170, loss: 8.329687261721119e-05
step: 180, loss: 0.00018651440041139722
step: 190, loss: 0.000190298305824399
step: 200, loss: 3.116479638265446e-05
step: 210, loss: 6.193658919073641e-05
step: 220, loss: 4.6521661715814844e-05
step: 230, loss: 0.00016051175771281123
/usr/local/lib/python3.7/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=0.9765363128491621, f1=0.977728285077951,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018988600641023368
step: 10, loss: 0.00031651632161810994
step: 20, loss: 0.0025193304754793644
step: 30, loss: 0.00016825592319946736
step: 40, loss: 0.00010818723239935935
step: 50, loss: 4.988238652003929e-05
step: 60, loss: 0.000404448714107275
step: 70, loss: 3.5172550269635394e-05
step: 80, loss: 5.771326323156245e-05
step: 90, loss: 4.408655877341516e-05
step: 100, loss: 4.472574801184237e-05
step: 110, loss: 0.014342390932142735
step: 120, loss: 0.00022486248053610325
step: 130, loss: 5.048907769378275e-05
step: 140, loss: 3.543753336998634e-05
step: 150, loss: 3.8164391298778355e-05
step: 160, loss: 8.172391972038895e-05
step: 170, loss: 6.562525959452614e-05
step: 180, loss: 4.742372766486369e-05

```

```

step: 190, loss: 0.0006974944844841957
step: 200, loss: 0.0005744736990891397
step: 210, loss: 2.653845695022028e-05
step: 220, loss: 6.0619011492235586e-05
step: 230, loss: 3.6338788049761206e-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=0.9765363128491621, f1=0.9755555555555556,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.407795520615764e-05
step: 10, loss: 0.001464334549382329
step: 20, loss: 3.921422830899246e-05
step: 30, loss: 3.8551719626411796e-05
step: 40, loss: 0.02265482395887375
step: 50, loss: 5.515987504622899e-05
step: 60, loss: 0.00010535735782468691
step: 70, loss: 5.580452852882445e-05
step: 80, loss: 0.00011145668395329267
step: 90, loss: 7.1633534389548e-05
step: 100, loss: 4.155731949140318e-05
step: 110, loss: 0.00015889416681602597
step: 120, loss: 2.3532116756541654e-05
step: 130, loss: 3.0651030101580545e-05
step: 140, loss: 4.31146145274397e-05
step: 150, loss: 0.00013450399274006486
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 160, loss: 0.0024484579917043447
step: 170, loss: 5.105702075525187e-05
step: 180, loss: 9.469208453083411e-05
step: 190, loss: 0.0012054521357640624
step: 200, loss: 5.295656228554435e-05
step: 210, loss: 4.8377009079558775e-05
step: 220, loss: 3.374631342012435e-05
step: 230, loss: 0.000719711824785918
/usr/local/lib/python3.7/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=0.9765363128491621, f1=0.9810055865921787,
best_f1=0.9755011135857461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.755207944777794e-05
step: 10, loss: 2.0999104890506715e-05

```

```

step: 20, loss: 0.00025159993674606085
step: 30, loss: 0.00012576566950883716
step: 40, loss: 0.00027998481527902186
step: 50, loss: 0.005368425976485014
step: 60, loss: 0.00010171304893447086
step: 70, loss: 6.76778145134449e-05
step: 80, loss: 7.416880544042215e-05
step: 90, loss: 8.682261977810413e-05
step: 100, loss: 0.0008303593494929373
step: 110, loss: 7.441522757289931e-05
step: 120, loss: 9.091122046811506e-05
step: 130, loss: 3.8294485420919955e-05
step: 140, loss: 4.842838097829372e-05
step: 150, loss: 0.00033934460952878
step: 160, loss: 3.228205241612159e-05
step: 170, loss: 2.703748941712547e-05
step: 180, loss: 0.0003549768007360399
step: 190, loss: 5.751214484917e-05
step: 200, loss: 5.563685408560559e-05
step: 210, loss: 0.00032665932667441666
step: 220, loss: 0.000331385585013777
step: 230, loss: 4.4147564040031284e-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=0.9775280898876404, f1=0.9808773903262092,
best_f1=0.9755011135857461
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 210.71it/s]
load_f1 = 0.9787709497206705
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 [00:17, 253.21it/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 bert-base-uncased \
    --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 bert-base-uncased \
    --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: 0.6307729482650757
step: 10, loss: 0.5116791129112244
step: 20, loss: 0.5376427173614502
step: 30, loss: 0.2225271612405777
step: 40, loss: 0.20386184751987457
step: 50, loss: 0.28114327788352966
step: 60, loss: 0.07726120203733444
step: 70, loss: 0.15436582267284393
step: 80, loss: 0.07224278897047043
step: 90, loss: 0.04765675961971283
step: 100, loss: 0.036880847066640854
step: 110, loss: 0.049555402249097824
```

```

step: 120, loss: 0.1729259192943573
step: 130, loss: 0.09313459694385529
step: 140, loss: 0.15318775177001953
step: 150, loss: 0.045988183468580246
step: 160, loss: 0.02520507574081421
step: 170, loss: 0.23363670706748962
step: 180, loss: 0.2048174887895584
step: 190, loss: 0.009783057495951653
step: 200, loss: 0.13948902487754822
step: 210, loss: 0.014135531149804592
step: 220, loss: 0.21009758114814758
step: 230, loss: 0.11048240214586258
step: 240, loss: 0.09313053637742996
step: 250, loss: 0.018163319677114487
step: 260, loss: 0.1410256177186966
step: 270, loss: 0.008299284614622593
step: 280, loss: 0.02369680255651474
step: 290, loss: 0.16910487413406372
step: 300, loss: 0.03308264538645744
step: 310, loss: 0.17405621707439423
step: 320, loss: 0.13624723255634308
step: 330, loss: 0.04177641496062279
step: 340, loss: 0.019859984517097473
step: 350, loss: 0.06292828917503357
step: 360, loss: 0.02557295747101307
step: 370, loss: 0.04449380189180374
step: 380, loss: 0.024509411305189133
step: 390, loss: 0.19033773243427277
step: 400, loss: 0.22776901721954346
step: 410, loss: 0.07091238349676132
step: 420, loss: 0.06574337184429169
step: 430, loss: 0.14066411554813385
step: 440, loss: 0.025944625958800316
step: 450, loss: 0.006461089476943016
step: 460, loss: 0.06049305200576782
step: 470, loss: 0.110963374376297
step: 480, loss: 0.04995698481798172
step: 490, loss: 0.22725415229797363
step: 500, loss: 0.12741875648498535
step: 510, loss: 0.05776817351579666
step: 520, loss: 0.12494073808193207
step: 530, loss: 0.004357177298516035
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9307479224376731, f1=0.9322820037105751,
best_f1=0.9322820037105751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1956339329481125
step: 10, loss: 0.05569889396429062
step: 20, loss: 0.018854748457670212
step: 30, loss: 0.05282512307167053
step: 40, loss: 0.21686111390590668
step: 50, loss: 0.1638326197862625
step: 60, loss: 0.008924935013055801
step: 70, loss: 0.027572283521294594
step: 80, loss: 0.04969339072704315
step: 90, loss: 0.024523308500647545
step: 100, loss: 0.013679927214980125
step: 110, loss: 0.008336649276316166
step: 120, loss: 0.1148570105433464
step: 130, loss: 0.21637113392353058
step: 140, loss: 0.017508935183286667
step: 150, loss: 0.052965156733989716
step: 160, loss: 0.030412796884775162
step: 170, loss: 0.02481074631214142
step: 180, loss: 0.06470821052789688
step: 190, loss: 0.08755320310592651
step: 200, loss: 0.004415736999362707
step: 210, loss: 0.015501068904995918
step: 220, loss: 0.05861566215753555
step: 230, loss: 0.002212373074144125
step: 240, loss: 0.023143306374549866
step: 250, loss: 0.07691628485918045
step: 260, loss: 0.002741017146036029
step: 270, loss: 0.3594549298286438
step: 280, loss: 0.056399811059236526
step: 290, loss: 0.025861982256174088
step: 300, loss: 0.18497507274150848
step: 310, loss: 0.04451332241296768
step: 320, loss: 0.03668956458568573
step: 330, loss: 0.025185072794556618
step: 340, loss: 0.03676968067884445
step: 350, loss: 0.0032595591619610786
step: 360, loss: 0.12251301109790802
step: 370, loss: 0.21622014045715332
step: 380, loss: 0.06104037165641785
step: 390, loss: 0.06023208051919937
step: 400, loss: 0.026824550703167915
step: 410, loss: 0.040040142834186554
```

```

step: 420, loss: 0.18377253413200378
step: 430, loss: 0.02355162799358368
step: 440, loss: 0.1538930982351303
step: 450, loss: 0.00979082752019167
step: 460, loss: 0.10711950808763504
step: 470, loss: 0.1068413034081459
step: 480, loss: 0.2521912157535553
step: 490, loss: 0.03982939198613167
step: 500, loss: 0.1801994889974594
step: 510, loss: 0.010388778522610664
step: 520, loss: 0.10765866190195084
step: 530, loss: 0.018529901280999184
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9340245051837888, f1=0.9316360207449317,
best_f1=0.9316360207449317
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.029775364324450493
step: 10, loss: 0.028834201395511627
step: 20, loss: 0.031453195959329605
step: 30, loss: 0.015082649886608124
step: 40, loss: 0.007488580886274576
step: 50, loss: 0.09476970881223679
step: 60, loss: 0.006936758290976286
step: 70, loss: 0.001819196972064674
step: 80, loss: 0.0014931465266272426
step: 90, loss: 0.007065081503242254
step: 100, loss: 0.0387478843331337
step: 110, loss: 0.0034156087785959244
step: 120, loss: 0.007899652235209942
step: 130, loss: 0.014267493970692158
step: 140, loss: 0.0029627024196088314
step: 150, loss: 0.04744500666856766
step: 160, loss: 0.022084074094891548
step: 170, loss: 0.08321832120418549
step: 180, loss: 0.026838865131139755
step: 190, loss: 0.00465679494664073
step: 200, loss: 0.10299994051456451
step: 210, loss: 0.04167623072862625
step: 220, loss: 0.027236230671405792
step: 230, loss: 0.02235463447868824

```

```

step: 240, loss: 0.0017375557217746973
step: 250, loss: 0.07403434813022614
step: 260, loss: 0.020068775862455368
step: 270, loss: 0.007868170738220215
step: 280, loss: 0.08289359509944916
step: 290, loss: 0.001590369618497789
step: 300, loss: 0.013612971641123295
step: 310, loss: 0.0032947801519185305
step: 320, loss: 0.07334110885858536
step: 330, loss: 0.0042465864680707455
step: 340, loss: 0.0052255019545555115
step: 350, loss: 0.02037326991558075
step: 360, loss: 0.05066382884979248
step: 370, loss: 0.005190294235944748
step: 380, loss: 0.03956708684563637
step: 390, loss: 0.02733069285750389
step: 400, loss: 0.017077479511499405
step: 410, loss: 0.002491460181772709
step: 420, loss: 0.05378052964806557
step: 430, loss: 0.005687069147825241
step: 440, loss: 0.010639277286827564
step: 450, loss: 0.08338530361652374
step: 460, loss: 0.04191947355866432
step: 470, loss: 0.015210014767944813
step: 480, loss: 0.0035969321615993977
step: 490, loss: 0.034034617245197296
step: 500, loss: 0.02031429298222065
step: 510, loss: 0.013829110190272331
step: 520, loss: 0.16164766252040863
step: 530, loss: 0.12363824993371964
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9344490934449092, f1=0.9278642149929278,
best_f1=0.9278642149929278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025446906685829163
step: 10, loss: 0.0034773030783981085
step: 20, loss: 0.004976049065589905
step: 30, loss: 0.002295037033036351
step: 40, loss: 0.01177167147397995
step: 50, loss: 0.001592107000760734

```

step: 60, loss: 0.0022238451056182384
step: 70, loss: 0.0007301029399968684
step: 80, loss: 0.023928072303533554
step: 90, loss: 0.017999669536948204
step: 100, loss: 0.015635516494512558
step: 110, loss: 0.016163045540452003
step: 120, loss: 0.00526600843295455
step: 130, loss: 0.002318625571206212
step: 140, loss: 0.002793633146211505
step: 150, loss: 0.02687184140086174
step: 160, loss: 0.0070847938768565655
step: 170, loss: 0.006022121757268906
step: 180, loss: 0.002460871357470751
step: 190, loss: 0.17507243156433105
step: 200, loss: 0.001833907444961369
step: 210, loss: 0.037873730063438416
step: 220, loss: 0.0010107691632583737
step: 230, loss: 0.009784498251974583
step: 240, loss: 0.003431156277656555
step: 250, loss: 0.008718075230717659
step: 260, loss: 0.0018110109958797693
step: 270, loss: 0.007736311759799719
step: 280, loss: 0.07104682922363281
step: 290, loss: 0.0526873916387558
step: 300, loss: 0.0006233923486433923
step: 310, loss: 0.06339360773563385
step: 320, loss: 0.03729701414704323
step: 330, loss: 0.0073302993550896645
step: 340, loss: 0.0036254168953746557
step: 350, loss: 0.025122124701738358
step: 360, loss: 0.029902102425694466
step: 370, loss: 0.003875532653182745
step: 380, loss: 0.05260513722896576
step: 390, loss: 0.0005119703128002584
step: 400, loss: 0.00099133828189224
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 410, loss: 0.009816375561058521
step: 420, loss: 0.0010612853802740574
step: 430, loss: 0.09651830792427063
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 440, loss: 0.0006519126472994685
step: 450, loss: 0.031518466770648956
step: 460, loss: 0.03149603679776192
step: 470, loss: 0.06135253980755806
step: 480, loss: 0.01583290658891201
step: 490, loss: 0.010550244711339474
step: 500, loss: 0.010369845665991306
step: 510, loss: 0.004077381454408169

```

step: 520, loss: 0.009896657429635525
step: 530, loss: 0.012936703860759735
/usr/local/lib/python3.7/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=0.9246823956442831, f1=0.9229374433363553,
best_f1=0.9278642149929278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1770845651626587
step: 10, loss: 0.025627033784985542
step: 20, loss: 0.07036726921796799
step: 30, loss: 0.0009378149406984448
step: 40, loss: 0.008883361704647541
step: 50, loss: 0.00034383885213173926
step: 60, loss: 0.0019190324237570167
step: 70, loss: 0.001175232115201652
step: 80, loss: 0.00019710586639121175
step: 90, loss: 0.011595739051699638
step: 100, loss: 0.0033333541359752417
step: 110, loss: 0.0010935108875855803
step: 120, loss: 0.0037058107554912567
step: 130, loss: 0.0015794640639796853
step: 140, loss: 0.0010975380428135395
step: 150, loss: 0.0006514160195365548
step: 160, loss: 0.00033538261777721345
step: 170, loss: 0.027575943619012833
step: 180, loss: 0.0011541841086000204
step: 190, loss: 0.002087540691718459
step: 200, loss: 0.0007022308418527246
step: 210, loss: 0.04400760307908058
step: 220, loss: 0.0023507673759013414
step: 230, loss: 0.023345032706856728
step: 240, loss: 0.004242179915308952
step: 250, loss: 0.006210905034095049
step: 260, loss: 0.045130763202905655
step: 270, loss: 0.0005774853634648025
step: 280, loss: 0.003703601658344269
step: 290, loss: 0.08529368042945862
step: 300, loss: 0.020729145035147667
step: 310, loss: 0.005008715204894543
step: 320, loss: 0.09313465654850006
step: 330, loss: 0.2164238542318344
step: 340, loss: 0.003742827335372567
step: 350, loss: 0.00026872940361499786
step: 360, loss: 0.0031054699793457985
step: 370, loss: 0.0032636860851198435

```

```

step: 380, loss: 0.0021992006804794073
step: 390, loss: 0.00011899645323865116
step: 400, loss: 0.0037751246709376574
step: 410, loss: 0.01963956654071808
step: 420, loss: 0.001938435947522521
step: 430, loss: 0.009836063720285892
step: 440, loss: 0.018099313601851463
step: 450, loss: 0.002753540640696883
step: 460, loss: 0.037813298404216766
step: 470, loss: 0.047347813844680786
step: 480, loss: 0.0647284984588623
step: 490, loss: 0.00023391087597701699
step: 500, loss: 0.0064266640692949295
step: 510, loss: 0.0028315503150224686
step: 520, loss: 0.0006852172664366663
step: 530, loss: 0.0033944682218134403
/usr/local/lib/python3.7/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=0.9289772727272728, f1=0.9221957040572791,
best_f1=0.9278642149929278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007899457705207169
step: 10, loss: 0.0018356702057644725
step: 20, loss: 0.07526040822267532
step: 30, loss: 0.001271457877010107
step: 40, loss: 0.0012050011428073049
step: 50, loss: 0.029327359050512314
step: 60, loss: 0.00012865409371443093
step: 70, loss: 0.0005679072928614914
step: 80, loss: 0.004402863793075085
step: 90, loss: 0.0005045935395173728
step: 100, loss: 0.06424573063850403
step: 110, loss: 0.0023204535245895386
step: 120, loss: 0.001719460473395884
step: 130, loss: 0.0005792190786451101
step: 140, loss: 0.0018747327849268913
step: 150, loss: 0.015484537929296494
step: 160, loss: 0.0008379866485483944
step: 170, loss: 0.001902903662994504
step: 180, loss: 0.000547240546438843
step: 190, loss: 0.00024769920855760574
step: 200, loss: 0.005001958925276995
step: 210, loss: 0.004075324162840843
step: 220, loss: 0.0007850155234336853
step: 230, loss: 0.00034833510289900005

```



```

step: 240, loss: 0.029256654903292656
step: 250, loss: 0.0003252677561249584
step: 260, loss: 0.000594722107052803
step: 270, loss: 0.17969726026058197
step: 280, loss: 0.002310763578861952
step: 290, loss: 0.0006011488731019199
step: 300, loss: 0.0024828847963362932
step: 310, loss: 0.0036665424704551697
step: 320, loss: 0.002756949979811907
step: 330, loss: 0.0018946018535643816
step: 340, loss: 0.0388069786131382
step: 350, loss: 0.0007401690818369389
step: 360, loss: 0.004737881012260914
step: 370, loss: 0.007484528701752424
step: 380, loss: 0.00019527898984961212
step: 390, loss: 0.0012342870468273759
step: 400, loss: 0.09845340251922607
step: 410, loss: 0.001962654758244753
step: 420, loss: 0.00164921919349581
step: 430, loss: 0.0020695501007139683
step: 440, loss: 0.0002024698187597096
step: 450, loss: 0.0022931837011128664
step: 460, loss: 0.000529334181919694
step: 470, loss: 0.006963694468140602
step: 480, loss: 0.0004548417928162962
step: 490, loss: 0.0011620430741459131
step: 500, loss: 0.002469703322276473
step: 510, loss: 0.0015499430010095239
step: 520, loss: 0.02110685035586357
step: 530, loss: 0.013853080570697784
/usr/local/lib/python3.7/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=0.9317647058823528, f1=0.9272388059701492,
best_f1=0.9278642149929278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008518265676684678
step: 10, loss: 0.0003188895061612129
step: 20, loss: 0.0026177912950515747
step: 30, loss: 0.0009447474731132388
step: 40, loss: 0.14101262390613556
step: 50, loss: 0.006388816516846418
step: 60, loss: 0.0007804523920640349
step: 70, loss: 0.0015211841091513634
step: 80, loss: 0.0007228325121104717
step: 90, loss: 0.0005229872767813504

```

```
step: 100, loss: 0.00039039907278493047
step: 110, loss: 0.004269953817129135
step: 120, loss: 5.672975385095924e-05
step: 130, loss: 0.002223587827757001
step: 140, loss: 0.009562644176185131
step: 150, loss: 0.00025823930627666414
step: 160, loss: 0.00012100509775336832
step: 170, loss: 0.0001103436152334325
step: 180, loss: 0.000130826752865687
step: 190, loss: 0.0008795917965471745
step: 200, loss: 0.00031367811607196927
step: 210, loss: 0.0014301524497568607
step: 220, loss: 0.00011602530867094174
step: 230, loss: 0.0015497241402044892
step: 240, loss: 0.0009624743834137917
step: 250, loss: 0.0005483580171130598
step: 260, loss: 0.0007152835023589432
step: 270, loss: 0.000675368937663734
step: 280, loss: 0.0008957774261943996
step: 290, loss: 0.0011441516689956188
step: 300, loss: 0.02453889325261116
step: 310, loss: 0.0007801610045135021
step: 320, loss: 0.0013250806368887424
step: 330, loss: 0.00764043815433979
step: 340, loss: 0.07312183082103729
step: 350, loss: 0.0005495835794135928
step: 360, loss: 0.08074939250946045
step: 370, loss: 0.0014101593988016248
step: 380, loss: 0.002382292877882719
step: 390, loss: 0.012848705984652042
step: 400, loss: 0.016806039959192276
step: 410, loss: 0.0018978673033416271
step: 420, loss: 0.0003778671962209046
step: 430, loss: 0.00447367038577795
step: 440, loss: 0.0005564495804719627
step: 450, loss: 0.0018483992898836732
step: 460, loss: 0.00011390688450774178
step: 470, loss: 0.002887047128751874
step: 480, loss: 0.006683684419840574
step: 490, loss: 0.029385792091488838
step: 500, loss: 0.00039817712968215346
step: 510, loss: 0.0012541778851300478
step: 520, loss: 0.022135842591524124
step: 530, loss: 0.0020877665374428034
/usr/local/lib/python3.7/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=0.933271547729379, f1=0.9314942528735632,
```

```
best_f1=0.9278642149929278
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02819335274398327
step: 10, loss: 0.00921015813946724
step: 20, loss: 0.00010754097456810996
step: 30, loss: 0.0023496488574892282
step: 40, loss: 8.397880446864292e-05
step: 50, loss: 0.0013359682634472847
step: 60, loss: 0.00670336838811636
step: 70, loss: 0.00011419531074352562
step: 80, loss: 0.000854643527418375
step: 90, loss: 0.00023014402540866286
step: 100, loss: 5.600736039923504e-05
step: 110, loss: 0.00037078381865285337
step: 120, loss: 0.0001882117794593796
step: 130, loss: 0.0015722776297479868
step: 140, loss: 0.0008303174399770796
step: 150, loss: 0.00020872097229585052
step: 160, loss: 0.000175420573214069
step: 170, loss: 0.08047032356262207
step: 180, loss: 6.60930818412453e-05
step: 190, loss: 0.00033548372448422015
step: 200, loss: 0.00015680260548833758
step: 210, loss: 0.00010210260370513424
step: 220, loss: 0.00011879372323164716
step: 230, loss: 4.6413493691943586e-05
step: 240, loss: 0.014237958006560802
step: 250, loss: 0.00013228955504018813
step: 260, loss: 0.00018935339176096022
step: 270, loss: 0.020259737968444824
step: 280, loss: 0.011237739585340023
step: 290, loss: 0.00018161864136345685
step: 300, loss: 0.0002089074841933325
step: 310, loss: 0.013971388339996338
step: 320, loss: 0.0018894897075369954
step: 330, loss: 0.00010498808114789426
step: 340, loss: 0.00016658162348903716
step: 350, loss: 0.00022842099133413285
step: 360, loss: 0.00016613276966381818
step: 370, loss: 4.042047658003867e-05
step: 380, loss: 0.0005590923829004169
step: 390, loss: 0.22995564341545105
step: 400, loss: 0.00022745836758986115
step: 410, loss: 0.0005369687569327652
step: 420, loss: 0.0011793707963079214
step: 430, loss: 0.003494466422125697
```

```

step: 440, loss: 0.0009923817124217749
step: 450, loss: 0.0001133238329202868
step: 460, loss: 6.975134601816535e-05
step: 470, loss: 0.00028412960818968713
step: 480, loss: 6.597091851290315e-05
step: 490, loss: 0.16673380136489868
step: 500, loss: 0.0037178401835262775
step: 510, loss: 0.002176918089389801
step: 520, loss: 0.007154872175306082
step: 530, loss: 0.00027233900618739426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9354988399071927, f1=0.9323447636700649,
best_f1=0.9323447636700649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011794296093285084
step: 10, loss: 5.05396346852649e-05
step: 20, loss: 0.010764137841761112
step: 30, loss: 0.001084189279936254
step: 40, loss: 4.7878125769784674e-05
step: 50, loss: 0.000449511397164315
step: 60, loss: 5.519023761735298e-05
step: 70, loss: 0.0011251728283241391
step: 80, loss: 0.002999397926032543
step: 90, loss: 0.00046814512461423874
step: 100, loss: 8.853139297571033e-05
step: 110, loss: 0.00012606017116922885
step: 120, loss: 0.00021971823298372328
step: 130, loss: 0.0004398797173053026
step: 140, loss: 0.0001219765908899717
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.00032053174800239503
step: 160, loss: 0.0110640162602067
step: 170, loss: 0.012584148906171322
step: 180, loss: 6.0380516515579075e-05
step: 190, loss: 7.734390237601474e-05
step: 200, loss: 8.477398660033941e-05
step: 210, loss: 2.751038846326992e-05
step: 220, loss: 9.316517389379442e-05
step: 230, loss: 3.405856841709465e-05
step: 240, loss: 0.0001894822926260531

```

```

step: 250, loss: 0.017288431525230408
step: 260, loss: 0.0009583849459886551
step: 270, loss: 9.02475367183797e-05
step: 280, loss: 0.00024916199618019164
step: 290, loss: 0.009605584666132927
step: 300, loss: 3.022590681212023e-05
step: 310, loss: 0.00036428726161830127
step: 320, loss: 0.03768037632107735
step: 330, loss: 6.540554750245064e-05
step: 340, loss: 0.00021719495998695493
step: 350, loss: 0.00028201600071042776
step: 360, loss: 3.837106123683043e-05
step: 370, loss: 0.008725748397409916
step: 380, loss: 0.00015865886234678328
step: 390, loss: 0.007047031540423632
step: 400, loss: 0.0006710326415486634
step: 410, loss: 4.742529563372955e-05
step: 420, loss: 0.0007305934559553862
step: 430, loss: 0.00019007727678399533
step: 440, loss: 0.0013610003516077995
step: 450, loss: 6.0645532357739285e-05
step: 460, loss: 0.00036710576387122273
step: 470, loss: 3.962566188420169e-05
step: 480, loss: 0.000815010629594326
step: 490, loss: 0.00023561548732686788
step: 500, loss: 0.001152914366684854
step: 510, loss: 0.00016412307741120458
step: 520, loss: 5.971561768092215e-05
step: 530, loss: 0.00010333766113035381
/usr/local/lib/python3.7/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=0.9351851851851852, f1=0.9317231769623782,
best_f1=0.9323447636700649
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00045040753320790827
step: 10, loss: 4.882032590103336e-05
step: 20, loss: 2.012351615121588e-05
step: 30, loss: 3.0531326046912e-05
step: 40, loss: 8.663292828714475e-05
step: 50, loss: 3.230785659980029e-05
step: 60, loss: 7.664111763006076e-05
step: 70, loss: 0.00012109129602322355
step: 80, loss: 2.3483646145905368e-05
step: 90, loss: 9.598115138942376e-05
step: 100, loss: 2.0347139070509e-05

```

```
step: 110, loss: 1.801889084163122e-05
step: 120, loss: 0.0001325858465861529
step: 130, loss: 0.00021922063024248928
step: 140, loss: 0.0006396176177076995
step: 150, loss: 0.0016113659366965294
step: 160, loss: 8.945615991251543e-05
step: 170, loss: 0.002193088410422206
step: 180, loss: 5.818776844535023e-05
step: 190, loss: 4.715450995718129e-05
step: 200, loss: 0.03992244228720665
step: 210, loss: 0.003543152706697583
step: 220, loss: 0.0011759598273783922
step: 230, loss: 0.001029586885124445
step: 240, loss: 0.01734039932489395
step: 250, loss: 3.491854658932425e-05
step: 260, loss: 0.0008076640078797936
step: 270, loss: 0.00015066830383148044
step: 280, loss: 0.0005538144032470882
step: 290, loss: 0.007172219455242157
step: 300, loss: 0.0003077126166317612
step: 310, loss: 7.263291627168655e-05
step: 320, loss: 0.0038240987341850996
step: 330, loss: 7.311560329981148e-05
step: 340, loss: 0.00013772047532256693
step: 350, loss: 7.59853792260401e-05
step: 360, loss: 5.759613850386813e-05
step: 370, loss: 0.024976136162877083
step: 380, loss: 0.0003714078338816762
step: 390, loss: 0.00011582868319237605
step: 400, loss: 0.00040050214738585055
step: 410, loss: 0.0007818436715751886
step: 420, loss: 0.00018921987793873996
step: 430, loss: 0.0001504958199802786
step: 440, loss: 0.00035373884020373225
step: 450, loss: 0.00192429986782372
step: 460, loss: 8.00935085862875e-05
step: 470, loss: 0.00023510016035288572
step: 480, loss: 0.00038721159216947854
step: 490, loss: 0.00028851875686086714
step: 500, loss: 0.013136486522853374
step: 510, loss: 0.0003679335059132427
step: 520, loss: 0.0004474966845009476
step: 530, loss: 0.0006667079287581146
/usr/local/lib/python3.7/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=0.9314685314685315, f1=0.9309056956115779,
best_f1=0.9323447636700649
```

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

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

```
step: 0, loss: 0.00012203966616652906  
step: 10, loss: 0.00010197170922765508  
step: 20, loss: 0.0003499736776575446  
step: 30, loss: 0.0003781066625379026  
step: 40, loss: 0.00020680668239947408  
step: 50, loss: 0.000877965590916574  
step: 60, loss: 0.00019878779130522162  
step: 70, loss: 0.00012980842438992113  
step: 80, loss: 2.376264637860004e-05  
step: 90, loss: 5.655417044181377e-05  
step: 100, loss: 0.0011397427879273891  
step: 110, loss: 6.330014730338007e-05  
step: 120, loss: 0.0022487116511911154  
step: 130, loss: 2.519971167203039e-05  
step: 140, loss: 0.00013433660205919296  
step: 150, loss: 1.9158618670189753e-05  
step: 160, loss: 3.197908154106699e-05  
step: 170, loss: 4.4718297431245446e-05  
step: 180, loss: 6.418812699848786e-05  
step: 190, loss: 0.0003919703303836286  
step: 200, loss: 0.0009653124725446105  
step: 210, loss: 4.9953105190070346e-05  
step: 220, loss: 7.826452929293737e-05  
step: 230, loss: 1.2174131370557006e-05  
step: 240, loss: 6.37678531347774e-05  
step: 250, loss: 1.4662481589766685e-05  
step: 260, loss: 2.2887430532136932e-05  
step: 270, loss: 7.505503162974492e-05  
step: 280, loss: 3.1689403840573505e-05  
step: 290, loss: 0.0009540563332848251  
step: 300, loss: 1.9482748029986396e-05  
step: 310, loss: 2.3401427824865095e-05  
step: 320, loss: 4.6731798647670075e-05  
step: 330, loss: 3.1305138691095635e-05  
step: 340, loss: 3.284812191850506e-05  
step: 350, loss: 0.0011516917729750276  
step: 360, loss: 1.92479037650628e-05  
step: 370, loss: 3.504452251945622e-05  
step: 380, loss: 2.1114477931405418e-05  
step: 390, loss: 1.73297303263098e-05  
step: 400, loss: 2.1468158593052067e-05  
step: 410, loss: 0.0012728903675451875  
step: 420, loss: 3.309731255285442e-05  
step: 430, loss: 9.097748989006504e-05  
step: 440, loss: 3.341200135764666e-05
```

```

step: 450, loss: 3.4290915209567174e-05
step: 460, loss: 0.027045512571930885
step: 470, loss: 3.288740117568523e-05
step: 480, loss: 0.00014010470476932824
step: 490, loss: 1.6651796613587067e-05
step: 500, loss: 4.47414604423102e-05
step: 510, loss: 3.343435673741624e-05
step: 520, loss: 4.869022450293414e-05
step: 530, loss: 0.002835912862792611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9359742054352833, f1=0.9354838709677419,
best_f1=0.9354838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.593479050323367e-05
step: 10, loss: 1.904706186905969e-05
step: 20, loss: 0.005009418353438377
step: 30, loss: 3.4545999369584024e-05
step: 40, loss: 0.00011539211118360981
step: 50, loss: 0.0035985936410725117
step: 60, loss: 0.001712589175440371
step: 70, loss: 4.228106263326481e-05
step: 80, loss: 0.00010726010077632964
step: 90, loss: 3.5661691072164103e-05
step: 100, loss: 0.00036333134630694985
step: 110, loss: 2.6761765184346586e-05
step: 120, loss: 4.925368193653412e-05
step: 130, loss: 0.0007238658727146685
step: 140, loss: 2.860103086277377e-05
step: 150, loss: 1.889796658360865e-05
step: 160, loss: 1.2654670172196347e-05
step: 170, loss: 5.051032349001616e-05
step: 180, loss: 2.357276389375329e-05
step: 190, loss: 5.038263043388724e-05
step: 200, loss: 2.027997834375128e-05
step: 210, loss: 0.0006255400949157774
step: 220, loss: 0.00013506195682566613
step: 230, loss: 4.6988159738248214e-05
step: 240, loss: 0.00012234764290042222
step: 250, loss: 8.568148041376844e-05
step: 260, loss: 2.7405802029534243e-05

```



```

step: 270, loss: 2.6784135116031393e-05
step: 280, loss: 0.0003642719238996506
step: 290, loss: 2.5491262931609526e-05
step: 300, loss: 0.005116807762533426
step: 310, loss: 7.478307088604197e-05
step: 320, loss: 2.7204023353988305e-05
step: 330, loss: 3.5756806028075516e-05
step: 340, loss: 5.160036016604863e-05
step: 350, loss: 0.0003018690913449973
step: 360, loss: 0.0014240018790587783
step: 370, loss: 6.595631566597149e-05
step: 380, loss: 5.339215204003267e-05
step: 390, loss: 0.00012374285142868757
step: 400, loss: 3.117927917628549e-05
step: 410, loss: 0.0045042079873383045
step: 420, loss: 0.01587119698524475
step: 430, loss: 0.0005446573486551642
step: 440, loss: 0.00015018266276456416
step: 450, loss: 0.0028901146724820137
step: 460, loss: 1.522130514786113e-05
step: 470, loss: 2.6071309548569843e-05
step: 480, loss: 6.50013898848556e-05
step: 490, loss: 0.0010173522168770432
step: 500, loss: 2.4734006728976965e-05
step: 510, loss: 0.00017343151557724923
step: 520, loss: 6.931564712431282e-05
step: 530, loss: 3.608088081819005e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9341486359360301, f1=0.922859830667921,
best_f1=0.9354838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019528381526470184
step: 10, loss: 2.105097155435942e-05
step: 20, loss: 6.889783253427595e-05
step: 30, loss: 3.65345949830953e-05
step: 40, loss: 2.3036251150188036e-05
step: 50, loss: 0.0006609720294363797
step: 60, loss: 1.9907409296138212e-05
step: 70, loss: 1.8678189007914625e-05
step: 80, loss: 1.2948968105774838e-05
step: 90, loss: 0.010899974964559078
step: 100, loss: 2.5732817448442802e-05
step: 110, loss: 1.6554815374547616e-05
step: 120, loss: 5.584873724728823e-05

```

```

step: 130, loss: 1.6309037164319307e-05
step: 140, loss: 5.861766112502664e-05
step: 150, loss: 2.469018909323495e-05
step: 160, loss: 2.5841345632215962e-05
step: 170, loss: 2.2276502932072617e-05
step: 180, loss: 7.951445149956271e-05
step: 190, loss: 1.5642277503502555e-05
step: 200, loss: 1.694607090030331e-05
step: 210, loss: 1.566824175824877e-05
step: 220, loss: 1.3015981494390871e-05
step: 230, loss: 1.5489462384721264e-05
step: 240, loss: 0.002197554800659418
step: 250, loss: 1.5444822565768845e-05
step: 260, loss: 1.6439233149867505e-05
step: 270, loss: 1.1358306437614374e-05
step: 280, loss: 1.9646458895294927e-05
step: 290, loss: 1.8949929653899744e-05
step: 300, loss: 5.940763367107138e-05
step: 310, loss: 2.2276572053669952e-05
step: 320, loss: 6.1005554016446695e-05
step: 330, loss: 3.8266651245066896e-05
step: 340, loss: 3.06825568259228e-05
step: 350, loss: 0.004696923308074474
step: 360, loss: 1.8674501916393638e-05
step: 370, loss: 2.3017910280032083e-05
step: 380, loss: 1.7363281585858203e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 390, loss: 0.0020458605140447617
step: 400, loss: 2.5292931240983307e-05
step: 410, loss: 0.0001926858676597476
step: 420, loss: 1.743404209264554e-05
step: 430, loss: 5.093269282951951e-05
step: 440, loss: 0.0001545355044072494
step: 450, loss: 2.1375015421654098e-05
step: 460, loss: 2.791984297800809e-05
step: 470, loss: 1.8885708414018154e-05
step: 480, loss: 1.4330988960864488e-05
step: 490, loss: 0.001458386075682938
step: 500, loss: 1.306435569858877e-05
step: 510, loss: 2.972466973005794e-05
step: 520, loss: 0.00036372223985381424
step: 530, loss: 1.9702010831679218e-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=0.9339578454332552, f1=0.9243776420854861,
best_f1=0.9354838709677419
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003139447420835495
step: 10, loss: 1.1097524293290917e-05
step: 20, loss: 1.3928666703577619e-05
step: 30, loss: 1.4547067621606402e-05
step: 40, loss: 3.7640678783645853e-05
step: 50, loss: 4.647160676540807e-05
step: 60, loss: 1.79290564119583e-05
step: 70, loss: 0.0001485950779169798
step: 80, loss: 2.6254376280121505e-05
step: 90, loss: 2.2644870114163496e-05
step: 100, loss: 0.00032973987981677055
step: 110, loss: 1.4677393664896954e-05
step: 120, loss: 7.221426494652405e-05
step: 130, loss: 0.005680023226886988
step: 140, loss: 3.9392161852447316e-05
step: 150, loss: 2.090810812660493e-05
step: 160, loss: 3.451183874858543e-05
step: 170, loss: 1.6860389223438688e-05
step: 180, loss: 1.6774714822531678e-05
step: 190, loss: 1.852528475865256e-05
step: 200, loss: 9.627616236684844e-05
step: 210, loss: 1.4535898117173929e-05
step: 220, loss: 2.0015231712022796e-05
step: 230, loss: 6.427543121390045e-05
step: 240, loss: 2.194105582020711e-05
step: 250, loss: 3.0148736186674796e-05
step: 260, loss: 8.888455340638757e-05
step: 270, loss: 1.361579415970482e-05
step: 280, loss: 1.3898868928663433e-05
step: 290, loss: 4.499302667682059e-05
step: 300, loss: 0.00017721332551445812
step: 310, loss: 0.0006483061006292701
step: 320, loss: 5.655493805534206e-05
step: 330, loss: 2.8828262657043524e-05
step: 340, loss: 4.0683255065232515e-05
step: 350, loss: 2.773676715150941e-05
step: 360, loss: 6.044032488716766e-05
step: 370, loss: 0.001720789703540504
step: 380, loss: 7.164126145653427e-05
step: 390, loss: 0.0013887012610211968
step: 400, loss: 0.0011823853710666299
step: 410, loss: 1.4740718143002596e-05
step: 420, loss: 2.6298712327843532e-05
step: 430, loss: 1.8775139324134216e-05
step: 440, loss: 4.751302913064137e-05
step: 450, loss: 5.3235708037391305e-05
```

```

step: 460, loss: 2.1311660020728596e-05
step: 470, loss: 0.00012996207806281745
step: 480, loss: 1.4226691746443976e-05
step: 490, loss: 0.00042156269773840904
step: 500, loss: 1.4308651770988945e-05
step: 510, loss: 0.00010535962064750493
step: 520, loss: 2.182919706683606e-05
step: 530, loss: 4.179935422143899e-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=0.9419414770088249, f1=0.9309225776541493,
best_f1=0.9309225776541493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.3965953257866204e-05
step: 10, loss: 0.0006544888601638377
step: 20, loss: 1.6830448657856323e-05
step: 30, loss: 2.4845530788297765e-05
step: 40, loss: 3.806559107033536e-05
step: 50, loss: 3.9495276723755524e-05
step: 60, loss: 1.3213392776378896e-05
step: 70, loss: 0.0003243098617531359
step: 80, loss: 6.641867366852239e-05
step: 90, loss: 1.9285429516457953e-05
step: 100, loss: 4.21962104155682e-05
step: 110, loss: 4.878882100456394e-05
step: 120, loss: 0.00015363722923211753
step: 130, loss: 0.07381158322095871
step: 140, loss: 1.657727079873439e-05
step: 150, loss: 2.0607712940545753e-05
step: 160, loss: 4.4449581764638424e-05
step: 170, loss: 5.41058907401748e-05
step: 180, loss: 1.1194388207513839e-05
step: 190, loss: 1.888319820864126e-05
step: 200, loss: 1.7247619325644337e-05
step: 210, loss: 0.00010816749272635207
step: 220, loss: 1.3161282367946114e-05
step: 230, loss: 3.768680107896216e-05
step: 240, loss: 1.491584498580778e-05
step: 250, loss: 1.9612971300375648e-05
step: 260, loss: 1.8510401787352748e-05
step: 270, loss: 1.6245732695097104e-05

```

```

step: 280, loss: 1.3772241800324991e-05
step: 290, loss: 1.1492402336443774e-05
step: 300, loss: 1.4442732208408415e-05
step: 310, loss: 4.887758768745698e-05
step: 320, loss: 1.546699604659807e-05
step: 330, loss: 1.8372484191786498e-05
step: 340, loss: 1.119812804972753e-05
step: 350, loss: 1.0512685548746958e-05
step: 360, loss: 0.000441272568423301
step: 370, loss: 1.148119099525502e-05
step: 380, loss: 3.0771981982979923e-05
step: 390, loss: 3.807251778198406e-05
step: 400, loss: 1.6826843420858495e-05
step: 410, loss: 0.00014150737843010575
step: 420, loss: 3.047790414711926e-05
step: 430, loss: 0.000406336213927716
step: 440, loss: 6.802468851674348e-05
step: 450, loss: 1.2077267456334084e-05
step: 460, loss: 1.4163354535412509e-05
step: 470, loss: 0.0001894988090498373
step: 480, loss: 1.0751075024018064e-05
step: 490, loss: 1.3749858226219658e-05
step: 500, loss: 1.7366832253173925e-05
step: 510, loss: 3.68289474863559e-05
step: 520, loss: 1.4155902135826182e-05
step: 530, loss: 1.8421036656945944e-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=0.9414519906323185, f1=0.9293785310734464,
best_f1=0.9309225776541493
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 [00:29, 194.24it/s]
load_f1 = 0.9410672853828307
real_f1 = 0.9410672853828307
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:22, 192.41it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                               | 827.9 MB 1.2 MB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: 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: numpy>=1.15.0 in /usr/local/lib/python3.7/dist-
packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.19.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.23.0)
Requirement already satisfied: 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: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.10.1)
Requirement already satisfied: srsly<1.1.0,>=0.1.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)
Requirement already satisfied: 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: 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: thinc<7.4.0,>=7.3.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)
Requirement already satisfied: blis<0.5.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)
Requirement already satisfied: catalogue<1.1.0,>=0.0.7 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.0)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.7)
Requirement already satisfied: importlib-metadata>=0.20 in
/usr/local/lib/python3.7/dist-packages (from
catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-
packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(3.8.1)
Requirement already satisfied: typing-extensions>=3.6.4 in
/usr/local/lib/python3.7/dist-packages (from importlib-
metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5)
(4.1.1)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)
Requirement already satisfied: certifi>=2017.4.17 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2022.6.15)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-
packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5)
(2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from
requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.0.4)
Requirement already satisfied: tqdm<5.0.0,>=4.10.0 in

```

```

/usr/local/lib/python3.7/dist-packages (from
thinc<7.4.0,>=7.3.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.41.0)
Building wheels for collected packages: en-core-web-lg
  Building wheel for en-core-web-lg (setup.py) ... done
  Created wheel for en-core-web-lg: filename=en_core_web_lg-2.2.5-py3-none-
any.whl size=829180942
sha256=94b98f69d9e610b44b6b9aa31bb203a2d6837a74388976384972c2a7dfb62195
  Stored in directory: /tmp/pip-ephem-wheel-cache-89n5fuka/wheels/11/95/ba/2c36c
c368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

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

```



```
--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>

rate

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: 0.5620827674865723

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

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

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

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=0.45161290322580644, f1=0.3829787234042553,
best_f1=0.3829787234042553
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4846442937850952
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6086956521739131, f1=0.4736842105263159,
best_f1=0.4736842105263159
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5578250288963318
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.4571428571428571,
best_f1=0.4571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2720649242401123
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.5000000000000001,
best_f1=0.4571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28687548637390137
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.4615384615384615,
best_f1=0.4571428571428571
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23377825319766998
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7586206896551724, f1=0.5555555555555556,
best_f1=0.5555555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09038093686103821
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787878, f1=0.5454545454545455,
best_f1=0.5454545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008914544247090816
/usr/local/lib/python3.7/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=0.7857142857142857, f1=0.6250000000000001,
best_f1=0.5454545454545455
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002683384343981743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.588235294117647,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014428733848035336
/usr/local/lib/python3.7/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=0.7741935483870968, f1=0.5714285714285714,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005240441299974918
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6285714285714286,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0054726446978747845
/usr/local/lib/python3.7/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=0.75, f1=0.6000000000000001, best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00855493638664484
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6000000000000001,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001994606340304017
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6000000000000001,
best_f1=0.588235294117647
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015845805406570435
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.6000000000000001,
best_f1=0.588235294117647
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 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, 120025.68it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7272727272727273
real_f1 = 0.6666666666666665
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 [00:18, 236.66it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.6059117317199707
step: 10, loss: 0.611181914806366
step: 20, loss: 0.33527982234954834
step: 30, loss: 0.1488988846540451
step: 40, loss: 0.3733484745025635
step: 50, loss: 0.11643394082784653
step: 60, loss: 0.16398367285728455
step: 70, loss: 0.02333519421517849
step: 80, loss: 0.020267415791749954
step: 90, loss: 0.3227471113204956
step: 100, loss: 0.019575025886297226
step: 110, loss: 0.15943145751953125
step: 120, loss: 0.006536425091326237
step: 130, loss: 0.003829285502433777
step: 140, loss: 0.0038007094990462065
step: 150, loss: 0.11099487543106079
step: 160, loss: 0.005907314829528332
step: 170, loss: 0.06413731724023819
step: 180, loss: 0.004143578466027975
step: 190, loss: 0.06911870092153549
step: 200, loss: 0.011374126188457012
step: 210, loss: 0.0057992469519376755
step: 220, loss: 0.004002366214990616
step: 230, loss: 0.08286087214946747
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
  warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9796839729119639, f1=0.9784335981838819,
best_f1=0.9784335981838819
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003988240379840136
step: 10, loss: 0.007441353984177113
step: 20, loss: 0.1463269740343094
step: 30, loss: 0.17317107319831848
step: 40, loss: 0.10396994650363922
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.009604064747691154
step: 60, loss: 0.0038037814665585756
step: 70, loss: 0.21021059155464172
step: 80, loss: 0.004350540228188038
step: 90, loss: 0.011855541728436947
step: 100, loss: 0.023767389357089996
step: 110, loss: 0.08303207904100418
step: 120, loss: 0.004572061821818352
step: 130, loss: 0.007278771605342627
step: 140, loss: 0.0024646075908094645
step: 150, loss: 0.0628281831741333
step: 160, loss: 0.025913668796420097
step: 170, loss: 0.0009532858384773135
step: 180, loss: 0.058723628520965576
step: 190, loss: 0.0017961438279598951
step: 200, loss: 0.0026363919023424387
step: 210, loss: 0.0013825084315612912
step: 220, loss: 0.1254175305366516
step: 230, loss: 0.039316289126873016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875424688561721, f1=0.9853438556933484,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.04970972612500191
step: 10, loss: 0.012848138809204102
step: 20, loss: 0.0012175787705928087
step: 30, loss: 0.003349388251081109
step: 40, loss: 0.026896502822637558
step: 50, loss: 0.01195070892572403
step: 60, loss: 0.0015384380239993334
step: 70, loss: 0.001769556780345738
step: 80, loss: 0.0006559052271768451
step: 90, loss: 0.19075515866279602
step: 100, loss: 0.0010318945860490203
step: 110, loss: 0.0008078203536570072
step: 120, loss: 0.046550195664167404
step: 130, loss: 0.0016172362957149744
step: 140, loss: 0.08001568913459778
step: 150, loss: 0.005265025421977043
step: 160, loss: 0.013753416016697884
step: 170, loss: 0.004979969467967749
step: 180, loss: 0.002972103189677
step: 190, loss: 0.0030539524741470814
step: 200, loss: 0.07336866110563278
step: 210, loss: 0.0023304440546780825
step: 220, loss: 0.0008848590077832341
step: 230, loss: 0.0036448698956519365
/usr/local/lib/python3.7/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=0.9821826280623607, f1=0.9788182831661093,
best_f1=0.9853438556933484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001709171338006854
step: 10, loss: 0.0003246757260058075
step: 20, loss: 0.00035910168662667274
step: 30, loss: 0.0003473885590210557
step: 40, loss: 0.03729569539427757
step: 50, loss: 0.0005903543205931783
step: 60, loss: 0.0038875846657902002
step: 70, loss: 0.0011626447085291147
step: 80, loss: 0.0051305051892995834
step: 90, loss: 0.0004760025185532868
step: 100, loss: 0.000518286949954927
step: 110, loss: 0.0006857928237877786
step: 120, loss: 0.07203167676925659
step: 130, loss: 0.001813827664591372
step: 140, loss: 0.0005909205647185445
step: 150, loss: 0.007443417329341173

```

```

step: 160, loss: 0.006245717406272888
step: 170, loss: 0.007255596108734608
step: 180, loss: 0.0003812692884821445
step: 190, loss: 0.0019119977951049805
step: 200, loss: 0.0015609340043738484
step: 210, loss: 0.05116428807377815
step: 220, loss: 0.0004838423046749085
step: 230, loss: 0.02201245352625847
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.990990990990991, 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: 0.0008141954895108938
step: 10, loss: 0.0006636914913542569
step: 20, loss: 0.0033733383752405643
step: 30, loss: 0.00018284606630913913
step: 40, loss: 0.0025158273056149483
step: 50, loss: 0.0003786934830714017
step: 60, loss: 0.0297209732234478
step: 70, loss: 0.0006489738589152694
step: 80, loss: 0.002148926490917802
step: 90, loss: 0.0023620750289410353
step: 100, loss: 0.0004112029855605215
step: 110, loss: 0.0002463505952619016
step: 120, loss: 0.0001233329821843654
step: 130, loss: 0.0014551416970789433
step: 140, loss: 0.005858397576957941
step: 150, loss: 0.00046098302118480206
step: 160, loss: 0.00035481381928548217
step: 170, loss: 0.008092770352959633
step: 180, loss: 0.015315856784582138
step: 190, loss: 0.09543842077255249
step: 200, loss: 0.02776476740837097
step: 210, loss: 0.0007673182990401983
step: 220, loss: 0.0011368737323209643
step: 230, loss: 0.0003057481371797621
/usr/local/lib/python3.7/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=0.9909502262443439, f1=0.9841628959276018,

```

```

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: 0.0005893422639928758
step: 10, loss: 0.0005842798855155706
step: 20, loss: 0.002685253042727709
step: 30, loss: 0.0005435132188722491
step: 40, loss: 0.00025670730974525213
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 50, loss: 0.00020609331841114908
step: 60, loss: 0.00015941115270834416
step: 70, loss: 0.020734811201691628
step: 80, loss: 0.0019428739324212074
step: 90, loss: 0.000609216804150492
step: 100, loss: 0.04423752427101135
step: 110, loss: 0.0003832468646578491
step: 120, loss: 0.00025113398442044854
step: 130, loss: 0.0011454051127657294
step: 140, loss: 0.00017147361359093338
step: 150, loss: 0.011167222633957863
step: 160, loss: 0.00021917707636021078
step: 170, loss: 0.00013773301907349378
step: 180, loss: 0.026753835380077362
step: 190, loss: 0.0002797854831442237
step: 200, loss: 0.0003968550590798259
step: 210, loss: 0.01548443827778101
step: 220, loss: 9.444540773984045e-05
step: 230, loss: 0.1650511920452118
/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.9819819819819819,
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: 0.0010745205217972398
step: 10, loss: 0.0001707050541881472
step: 20, loss: 0.006048272363841534
step: 30, loss: 0.00037031190004199743
step: 40, loss: 0.00036385853309184313
step: 50, loss: 0.00034542905632406473
step: 60, loss: 0.00021317480423022062
step: 70, loss: 0.014782070182263851
step: 80, loss: 0.00032083812402561307
step: 90, loss: 0.00014653820835519582
step: 100, loss: 0.0004341880267020315

```

```

step: 110, loss: 0.00010795980779221281
step: 120, loss: 8.580723078921437e-05
step: 130, loss: 0.00016396892897319049
step: 140, loss: 0.0007789547671563923
step: 150, loss: 0.014049806632101536
step: 160, loss: 0.04862329736351967
step: 170, loss: 0.00045240053441375494
step: 180, loss: 0.001250546076335013
step: 190, loss: 0.00015863694716244936
step: 200, loss: 0.08243868499994278
step: 210, loss: 5.639364826492965e-05
step: 220, loss: 0.0033671052660793066
step: 230, loss: 0.00012422299187164754
/usr/local/lib/python3.7/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=0.9898074745186863, f1=0.9829738933030647,
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: 8.274467836599797e-05
step: 10, loss: 0.00011986812751274556
step: 20, loss: 0.00011410965089453384
step: 30, loss: 0.00016938730550464243
step: 40, loss: 0.0002330265415366739
step: 50, loss: 0.0005765188834629953
step: 60, loss: 0.0001495241594966501
step: 70, loss: 0.00025532825384289026
step: 80, loss: 0.00021628881222568452
step: 90, loss: 4.124809129280038e-05
step: 100, loss: 0.00015629148401785642
step: 110, loss: 0.02799215540289879
step: 120, loss: 0.00022441506735049188
step: 130, loss: 0.006068977061659098
step: 140, loss: 7.979785732459277e-05
step: 150, loss: 0.0003779448161367327
step: 160, loss: 0.00010077144543174654
step: 170, loss: 5.3719308198196813e-05
step: 180, loss: 6.460365693783388e-05
step: 190, loss: 0.00011368478590156883
step: 200, loss: 5.719358887290582e-05
step: 210, loss: 6.769879109924659e-05
step: 220, loss: 0.00024374874192290008
step: 230, loss: 7.641538104508072e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

epoch 8: dev_f1=0.9887640449438202, f1=0.9808773903262092,
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: 6.53499664622359e-05
step: 10, loss: 0.00015581565094180405
step: 20, loss: 0.00021632153948303312
step: 30, loss: 4.9274993216386065e-05
step: 40, loss: 0.030559314414858818
step: 50, loss: 7.219148392323405e-05
step: 60, loss: 0.0002764719829428941
step: 70, loss: 0.025525685399770737
step: 80, loss: 0.0013233562931418419
step: 90, loss: 0.00018368272867519408
step: 100, loss: 0.007010356057435274
step: 110, loss: 5.4474712669616565e-05
step: 120, loss: 6.81498131598346e-05
step: 130, loss: 0.11074821650981903
step: 140, loss: 0.010614493861794472
step: 150, loss: 0.0014064532006159425
step: 160, loss: 0.00011158398410771042
step: 170, loss: 0.0002899171377066523
step: 180, loss: 0.002971445210278034
step: 190, loss: 0.00010279826528858393
step: 200, loss: 9.382270945934579e-05
step: 210, loss: 0.0003665972617454827
step: 220, loss: 6.749962631147355e-05
step: 230, loss: 9.279977530241013e-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=0.9854096520763187, f1=0.9796839729119639,
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: 0.00020971556659787893
step: 10, loss: 0.00012665620306506753
step: 20, loss: 9.000847057905048e-05
step: 30, loss: 0.00012679870997089893
step: 40, loss: 9.51701367739588e-05
step: 50, loss: 0.00035950797609984875
step: 60, loss: 0.0036881454288959503
step: 70, loss: 0.001616881461814046
step: 80, loss: 7.505427493015304e-05
step: 90, loss: 0.0016781880985945463
step: 100, loss: 0.0001268952473765239

```

```

step: 110, loss: 0.00010481038771104068
step: 120, loss: 9.25728163565509e-05
step: 130, loss: 6.985836807871237e-05
step: 140, loss: 0.05796099454164505
step: 150, loss: 0.008026476949453354
step: 160, loss: 3.6272067518439144e-05
step: 170, loss: 0.00010077381011797115
step: 180, loss: 9.308586595579982e-05
step: 190, loss: 0.01228022575378418
step: 200, loss: 0.00011065325088566169
step: 210, loss: 9.6011339337565e-05
step: 220, loss: 6.099817619542591e-05
step: 230, loss: 0.0001882334763649851
/usr/local/lib/python3.7/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=0.9843400447427293, f1=0.978675645342312,
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: 0.0001258408883586526
step: 10, loss: 0.01137105468660593
step: 20, loss: 0.002546819858253002
step: 30, loss: 0.020896660163998604
step: 40, loss: 0.0001037332767737098
step: 50, loss: 0.00010070372809423134
step: 60, loss: 7.822015322744846e-05
step: 70, loss: 0.0004056546895299107
step: 80, loss: 0.00010807038052007556
step: 90, loss: 3.595529778976925e-05
step: 100, loss: 4.429497494129464e-05
step: 110, loss: 0.029341477900743484
step: 120, loss: 4.278232154319994e-05
step: 130, loss: 6.635719182668254e-05
step: 140, loss: 5.0155365897808224e-05
step: 150, loss: 0.030899394303560257
step: 160, loss: 3.65476735169068e-05
step: 170, loss: 0.028475528582930565
step: 180, loss: 5.4377167543862015e-05
step: 190, loss: 5.5573491408722475e-05
step: 200, loss: 0.0011739627225324512
step: 210, loss: 3.838403426925652e-05
step: 220, loss: 3.6581277527147904e-05
step: 230, loss: 0.00015026355686131865
/usr/local/lib/python3.7/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=0.9887387387387387, f1=0.9807474518686297,
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: 4.226173405186273e-05
step: 10, loss: 3.5996228689327836e-05
step: 20, loss: 3.379099871381186e-05
step: 30, loss: 4.6184606617316604e-05
step: 40, loss: 0.0004856008163187653
step: 50, loss: 0.00010237043898086995
step: 60, loss: 0.00013591408787760884
step: 70, loss: 0.000504637835547328
step: 80, loss: 7.81681010266766e-05
step: 90, loss: 5.63038483960554e-05
step: 100, loss: 8.089582115644589e-05
step: 110, loss: 4.77645626233425e-05
step: 120, loss: 5.1924489525845274e-05
step: 130, loss: 3.676011328934692e-05
step: 140, loss: 0.012259208597242832
step: 150, loss: 0.0006830113707110286
step: 160, loss: 4.470878775464371e-05
step: 170, loss: 4.1107021388597786e-05
step: 180, loss: 0.02040359377861023
step: 190, loss: 5.1823139074258506e-05
step: 200, loss: 3.5275148547952995e-05
step: 210, loss: 5.8294401242164895e-05
step: 220, loss: 4.212279600324109e-05
step: 230, loss: 0.03963029012084007
/usr/local/lib/python3.7/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=0.9886877828054299, f1=0.9829738933030647,
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: 6.55151525279507e-05
step: 10, loss: 0.0008782543009147048
step: 20, loss: 0.0001243498845724389
step: 30, loss: 0.00030429053003899753
step: 40, loss: 8.296417217934504e-05
step: 50, loss: 8.671992691233754e-05
step: 60, loss: 5.97627367824316e-05
step: 70, loss: 2.6519788661971688e-05
step: 80, loss: 4.7082765377126634e-05
step: 90, loss: 7.379040471278131e-05
step: 100, loss: 3.2170111808227375e-05

```

```

step: 110, loss: 0.033507172018289566
step: 120, loss: 0.025960084050893784
step: 130, loss: 5.6332413805648685e-05
step: 140, loss: 3.760171966860071e-05
step: 150, loss: 4.605573121807538e-05
step: 160, loss: 0.00026769412215799093
step: 170, loss: 0.0003948290541302413
step: 180, loss: 5.1348310080356896e-05
step: 190, loss: 3.594799272832461e-05
step: 200, loss: 4.7451041609747335e-05
step: 210, loss: 4.792941399500705e-05
step: 220, loss: 0.00017468795704189688
step: 230, loss: 0.00013524311361834407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.9898534385569334, f1=0.9797752808988766,
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: 3.613052467699163e-05
step: 10, loss: 0.010059130378067493
step: 20, loss: 0.0008100527920760214
step: 30, loss: 0.00018335790082346648
step: 40, loss: 0.0002461541153024882
step: 50, loss: 4.7122579417191446e-05
step: 60, loss: 8.842068200465292e-05
step: 70, loss: 5.8242338127456605e-05
step: 80, loss: 4.604601417668164e-05
step: 90, loss: 8.223658369388431e-05
step: 100, loss: 0.00010739926801761612
step: 110, loss: 6.222187221283093e-05
step: 120, loss: 2.5647992515587248e-05
step: 130, loss: 5.602025703410618e-05
step: 140, loss: 7.963574898894876e-05
step: 150, loss: 4.72661922685802e-05
step: 160, loss: 0.009751527570188046
step: 170, loss: 4.3671032472047955e-05
step: 180, loss: 4.259903289494105e-05
step: 190, loss: 0.0001116966232075356
step: 200, loss: 4.605402864399366e-05
step: 210, loss: 7.239599653985351e-05
step: 220, loss: 0.00014574134547729045
step: 230, loss: 0.014314772561192513
/usr/local/lib/python3.7/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=0.9909502262443439, f1=0.9807909604519773,
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: 3.269200897193514e-05
step: 10, loss: 2.5823113901424222e-05
step: 20, loss: 4.82327195641119e-05
step: 30, loss: 6.0782491345889866e-05
step: 40, loss: 8.010570309124887e-05
step: 50, loss: 0.05141625553369522
step: 60, loss: 0.00010812245454872027
step: 70, loss: 7.4500021582935e-05
step: 80, loss: 4.281180736143142e-05
step: 90, loss: 4.8434474592795596e-05
step: 100, loss: 3.9720875065540895e-05
step: 110, loss: 0.00016169129230547696
step: 120, loss: 6.315074278973043e-05
step: 130, loss: 0.013024765066802502
step: 140, loss: 3.1321487767854705e-05
step: 150, loss: 0.0003681635425891727
step: 160, loss: 3.1239265808835626e-05
step: 170, loss: 2.832644895534031e-05
step: 180, loss: 0.00018100447778124362
step: 190, loss: 0.002522439928725362
step: 200, loss: 0.00011340812488924712
step: 210, loss: 4.514778993325308e-05
step: 220, loss: 3.93967711715959e-05
step: 230, loss: 2.987979314639233e-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=0.9909502262443439, f1=0.9807909604519773,
best_f1=0.9842342342342343
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 184.50it/s]
load_f1 = 0.9887892376681614
real_f1 = 0.9865771812080537
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 224.52it/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 bert-base-uncased \
    --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 bert-base-uncased \
  --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: 0.618728518486023
step: 10, loss: 0.515981912612915
step: 20, loss: 0.5366967916488647
step: 30, loss: 0.21571069955825806
step: 40, loss: 0.2407843917608261
step: 50, loss: 0.13492625951766968
step: 60, loss: 0.04225018247961998
step: 70, loss: 0.06414006650447845
step: 80, loss: 0.12169679999351501
step: 90, loss: 0.48164230585098267
step: 100, loss: 0.04567006975412369
step: 110, loss: 0.08598248660564423
step: 120, loss: 0.127407044172287
step: 130, loss: 0.11097394675016403
step: 140, loss: 0.18651317059993744
step: 150, loss: 0.07540961354970932
step: 160, loss: 0.022495217621326447
step: 170, loss: 0.3126569092273712
step: 180, loss: 0.04699435830116272
step: 190, loss: 0.010290445759892464
step: 200, loss: 0.1928737908601761
step: 210, loss: 0.07692752778530121
step: 220, loss: 0.2660588324069977
step: 230, loss: 0.14805693924427032
step: 240, loss: 0.023193135857582092
step: 250, loss: 0.029829569160938263
step: 260, loss: 0.16553150117397308
step: 270, loss: 0.005201591178774834
step: 280, loss: 0.040971092879772186
step: 290, loss: 0.17219054698944092
step: 300, loss: 0.06331434845924377
step: 310, loss: 0.2297128587961197
step: 320, loss: 0.045154646039009094
step: 330, loss: 0.07781630009412766
step: 340, loss: 0.07056573033332825
step: 350, loss: 0.10790145397186279
step: 360, loss: 0.027928847819566727
step: 370, loss: 0.055861204862594604
step: 380, loss: 0.012857172638177872
step: 390, loss: 0.22029952704906464
step: 400, loss: 0.36013442277908325
step: 410, loss: 0.09282002598047256
step: 420, loss: 0.05812486261129379
step: 430, loss: 0.2337077409029007
step: 440, loss: 0.009490055963397026
step: 450, loss: 0.022324757650494576
step: 460, loss: 0.009175543673336506

```

step: 470, loss: 0.16353251039981842
step: 480, loss: 0.08413874357938766
step: 490, loss: 0.052150908857584
step: 500, loss: 0.12615440785884857
step: 510, loss: 0.03333437442779541
step: 520, loss: 0.1266438513994217
step: 530, loss: 0.0035739082377403975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9317972350230416, f1=0.9323515876668201,
best_f1=0.9323515876668201
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06924884021282196
step: 10, loss: 0.0393051914870739
step: 20, loss: 0.036855049431324005
step: 30, loss: 0.033584114164114
step: 40, loss: 0.06314270943403244
step: 50, loss: 0.2168661504983902
step: 60, loss: 0.0333644300699234
step: 70, loss: 0.03298493102192879
step: 80, loss: 0.02077917754650116
step: 90, loss: 0.10082071274518967
step: 100, loss: 0.014607016928493977
step: 110, loss: 0.06842157244682312
step: 120, loss: 0.053143154829740524
step: 130, loss: 0.07677986472845078
step: 140, loss: 0.014691661112010479
step: 150, loss: 0.11031849682331085
step: 160, loss: 0.02405460551381111
step: 170, loss: 0.0505279116332531
step: 180, loss: 0.016228802502155304
step: 190, loss: 0.12627500295639038
step: 200, loss: 0.013167462311685085
step: 210, loss: 0.11771532148122787
step: 220, loss: 0.06581374257802963
step: 230, loss: 0.017326481640338898
step: 240, loss: 0.053097914904356
step: 250, loss: 0.06455294787883759
step: 260, loss: 0.014858880080282688
step: 270, loss: 0.2529503107070923
step: 280, loss: 0.04549906775355339

```

```

step: 290, loss: 0.01780070923268795
step: 300, loss: 0.17117686569690704
step: 310, loss: 0.005995849147439003
step: 320, loss: 0.08937830477952957
step: 330, loss: 0.017817141488194466
step: 340, loss: 0.037942223250865936
step: 350, loss: 0.004653653595596552
step: 360, loss: 0.023059561848640442
step: 370, loss: 0.17343777418136597
step: 380, loss: 0.05995073541998863
step: 390, loss: 0.11288518458604813
step: 400, loss: 0.08719757944345474
step: 410, loss: 0.07190967351198196
step: 420, loss: 0.01557641290128231
step: 430, loss: 0.010183734819293022
step: 440, loss: 0.06735852360725403
step: 450, loss: 0.01872827485203743
step: 460, loss: 0.11252516508102417
step: 470, loss: 0.01583481766283512
step: 480, loss: 0.19921506941318512
step: 490, loss: 0.02389358915388584
step: 500, loss: 0.1790066659450531
step: 510, loss: 0.05235062912106514
step: 520, loss: 0.04379424825310707
step: 530, loss: 0.025343814864754677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9345182413470534, f1=0.9272137227630968,
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03808059170842171
step: 10, loss: 0.012615501880645752
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.07439304143190384
step: 30, loss: 0.03520280495285988
step: 40, loss: 0.0064519718289375305
step: 50, loss: 0.07310600578784943
step: 60, loss: 0.010614286176860332
step: 70, loss: 0.004157107789069414
step: 80, loss: 0.0010762747842818499
step: 90, loss: 0.002474491950124502

```

```
step: 100, loss: 0.14437077939510345
step: 110, loss: 0.007583060767501593
step: 120, loss: 0.005649389233440161
step: 130, loss: 0.019664710387587547
step: 140, loss: 0.03150424361228943
step: 150, loss: 0.008362935855984688
step: 160, loss: 0.015260039828717709
step: 170, loss: 0.04539187625050545
step: 180, loss: 0.010830646380782127
step: 190, loss: 0.012899565510451794
step: 200, loss: 0.008999058045446873
step: 210, loss: 0.015198147855699062
step: 220, loss: 0.02604391612112522
step: 230, loss: 0.06317012012004852
step: 240, loss: 0.0015216391766443849
step: 250, loss: 0.006632383447140455
step: 260, loss: 0.01637723110616207
step: 270, loss: 0.011745711788535118
step: 280, loss: 0.2106398195028305
step: 290, loss: 0.003980844747275114
step: 300, loss: 0.08755360543727875
step: 310, loss: 0.016442691907286644
step: 320, loss: 0.009631489403545856
step: 330, loss: 0.0029171414207667112
step: 340, loss: 0.00916792917996645
step: 350, loss: 0.0032869947608560324
step: 360, loss: 0.06576641649007797
step: 370, loss: 0.011902594938874245
step: 380, loss: 0.01596098765730858
step: 390, loss: 0.04861151799559593
step: 400, loss: 0.01756276749074459
step: 410, loss: 0.008160505443811417
step: 420, loss: 0.12525708973407745
step: 430, loss: 0.003640108974650502
step: 440, loss: 0.0028777415864169598
step: 450, loss: 0.06696987897157669
step: 460, loss: 0.013824112713336945
step: 470, loss: 0.03724226355552673
step: 480, loss: 0.004180482123047113
step: 490, loss: 0.0032680071890354156
step: 500, loss: 0.11082301288843155
step: 510, loss: 0.042007558047771454
step: 520, loss: 0.035006940364837646
step: 530, loss: 0.11177602410316467
/usr/local/lib/python3.7/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=0.9327808471454879, f1=0.9217311233885819,
```

```
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010976339690387249
step: 10, loss: 0.006677596364170313
step: 20, loss: 0.0021080428268760443
step: 30, loss: 0.007920732721686363
step: 40, loss: 0.10763366520404816
step: 50, loss: 0.0201860424131155
step: 60, loss: 0.006347812246531248
step: 70, loss: 0.03340178728103638
step: 80, loss: 0.003394537139683962
step: 90, loss: 0.04741382226347923
step: 100, loss: 0.002090269699692726
step: 110, loss: 0.021722150966525078
step: 120, loss: 0.0005662609473802149
step: 130, loss: 0.023779893293976784
step: 140, loss: 0.09869711101055145
step: 150, loss: 0.00780758960172534
step: 160, loss: 0.06011227145791054
step: 170, loss: 0.004733314272016287
step: 180, loss: 0.029182568192481995
step: 190, loss: 0.02283570170402527
step: 200, loss: 0.007789176423102617
step: 210, loss: 0.018967753276228905
step: 220, loss: 0.015121948905289173
step: 230, loss: 0.01139984279870987
step: 240, loss: 0.005227621644735336
step: 250, loss: 0.14548179507255554
step: 260, loss: 0.0021902977023273706
step: 270, loss: 0.0380401611328125
step: 280, loss: 0.05910724401473999
step: 290, loss: 0.020545054227113724
step: 300, loss: 0.0024234207812696695
step: 310, loss: 0.041709404438734055
step: 320, loss: 0.0808263048529625
step: 330, loss: 0.045657578855752945
step: 340, loss: 0.005275183357298374
step: 350, loss: 0.0226899404078722
step: 360, loss: 0.05096912384033203
step: 370, loss: 0.0017130578635260463
step: 380, loss: 0.008446352556347847
step: 390, loss: 0.019888920709490776
step: 400, loss: 0.011690815910696983
step: 410, loss: 0.0028103606309741735
step: 420, loss: 0.01021298486739397
step: 430, loss: 0.18208596110343933
```



```

step: 440, loss: 0.0031709212344139814
step: 450, loss: 0.0015954433474689722
step: 460, loss: 0.0023771722335368395
step: 470, loss: 0.0149462278932333
step: 480, loss: 0.02222374640405178
step: 490, loss: 0.005124672316014767
step: 500, loss: 0.011940614320337772
step: 510, loss: 0.004713071510195732
step: 520, loss: 0.07507845759391785
step: 530, loss: 0.0027739207725971937
/usr/local/lib/python3.7/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=0.9314814814814815, f1=0.9271889400921659,
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011368962936103344
step: 10, loss: 0.03258872404694557
step: 20, loss: 0.010979508981108665
step: 30, loss: 0.008422796614468098
step: 40, loss: 0.06580749154090881
step: 50, loss: 0.013733502477407455
step: 60, loss: 0.13756795227527618
step: 70, loss: 0.00159462948795408
step: 80, loss: 0.0003334791399538517
step: 90, loss: 0.031070811673998833
step: 100, loss: 0.11044403910636902
step: 110, loss: 0.0017502681585028768
step: 120, loss: 0.013588091358542442
step: 130, loss: 0.0014546706806868315
step: 140, loss: 0.0004966126289218664
step: 150, loss: 0.0026454529725015163
step: 160, loss: 0.0003865393518935889
step: 170, loss: 0.10669172555208206
step: 180, loss: 0.002379161538556218
step: 190, loss: 0.007829147391021252
step: 200, loss: 0.002733712550252676
step: 210, loss: 0.0013530271826311946
step: 220, loss: 0.0012457524426281452
step: 230, loss: 0.0061980560421943665
step: 240, loss: 0.003778153331950307
step: 250, loss: 0.021101979538798332
step: 260, loss: 0.00194891809951514
step: 270, loss: 0.0014987303875386715
step: 280, loss: 0.003239057259634137
step: 290, loss: 0.12224928289651871

```

```

step: 300, loss: 0.004826186690479517
step: 310, loss: 0.010258536785840988
step: 320, loss: 0.18420174717903137
step: 330, loss: 0.011600874364376068
step: 340, loss: 0.0045065986923873425
step: 350, loss: 0.0009025849867612123
step: 360, loss: 0.012427293695509434
step: 370, loss: 0.04348345845937729
step: 380, loss: 0.003230612026527524
step: 390, loss: 0.00023009035794530064
step: 400, loss: 0.0017212703824043274
step: 410, loss: 0.0010567621793597937
step: 420, loss: 0.003495460608974099
step: 430, loss: 0.001293838256970048
step: 440, loss: 0.012865149416029453
step: 450, loss: 0.005949912127107382
step: 460, loss: 0.0008740106713958085
step: 470, loss: 0.0010083181550726295
step: 480, loss: 0.003879395080730319
step: 490, loss: 0.0006579341716133058
step: 500, loss: 0.005530611611902714
step: 510, loss: 0.15052059292793274
step: 520, loss: 0.010439681820571423
step: 530, loss: 0.013736148364841938
/usr/local/lib/python3.7/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=0.9233593391463975,
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00702802324667573
step: 10, loss: 0.00046530886902473867
step: 20, loss: 0.10398077219724655
step: 30, loss: 0.009261524304747581
step: 40, loss: 0.11607876420021057
step: 50, loss: 0.021127792075276375
step: 60, loss: 0.052770182490348816
step: 70, loss: 0.0008609380456618965
step: 80, loss: 0.001721388311125338
step: 90, loss: 0.0036636090371757746
step: 100, loss: 0.00036452373024076223
step: 110, loss: 0.000707692583091557
step: 120, loss: 0.003209654474630952
step: 130, loss: 0.00024813375785015523
step: 140, loss: 0.0061063747853040695
step: 150, loss: 0.0009065705817192793

```

```
step: 160, loss: 0.001689845696091652
step: 170, loss: 0.006255008280277252
step: 180, loss: 0.00031491630943492055
step: 190, loss: 0.0007160500390455127
step: 200, loss: 0.0024839637335389853
step: 210, loss: 0.012732835486531258
step: 220, loss: 0.0005652702529914677
step: 230, loss: 0.00045395284541882575
step: 240, loss: 0.0010749832727015018
step: 250, loss: 0.000698892108630389
step: 260, loss: 0.00011744433868443593
step: 270, loss: 0.011293946765363216
step: 280, loss: 0.0014424154069274664
step: 290, loss: 0.0017396352486684918
step: 300, loss: 0.0009681288502179086
step: 310, loss: 0.0004578132356982678
step: 320, loss: 0.00024547247448936105
step: 330, loss: 0.00023475805937778205
step: 340, loss: 0.2152595967054367
step: 350, loss: 0.0012219587806612253
step: 360, loss: 0.0052760932594537735
step: 370, loss: 0.0050185490399599075
step: 380, loss: 0.00204872852191329
step: 390, loss: 0.03831702843308449
step: 400, loss: 0.00029731480753980577
step: 410, loss: 0.004156865645200014
step: 420, loss: 0.0033292374573647976
step: 430, loss: 0.001921538612805307
step: 440, loss: 7.239123806357384e-05
step: 450, loss: 0.0022078058682382107
step: 460, loss: 0.0019408121006563306
step: 470, loss: 0.0001475482276873663
step: 480, loss: 0.003043810836970806
step: 490, loss: 0.006051321048289537
step: 500, loss: 0.000978070660494268
step: 510, loss: 0.01049328688532114
step: 520, loss: 0.006560326553881168
step: 530, loss: 0.04426606744527817
/usr/local/lib/python3.7/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=0.9292740046838408, f1=0.9197012138188609,
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016936320753302425
step: 10, loss: 0.0017768391408026218
```

step: 20, loss: 0.0011564097367227077
step: 30, loss: 0.0003067585057578981
step: 40, loss: 0.027487298473715782
step: 50, loss: 0.00012846343452110887
step: 60, loss: 0.0018352663610130548
step: 70, loss: 0.0010446715168654919
step: 80, loss: 0.0004950986476615071
step: 90, loss: 0.00042814022162929177
step: 100, loss: 0.00012238803901709616
step: 110, loss: 0.0036210371181368828
step: 120, loss: 0.00028730204212479293
step: 130, loss: 0.0003740586689673364
step: 140, loss: 0.00016972301818896085
step: 150, loss: 0.00023170662461780012
step: 160, loss: 7.270366768352687e-05
step: 170, loss: 0.0001479435886722058
step: 180, loss: 0.0005932722124271095
step: 190, loss: 0.001290147309191525
step: 200, loss: 0.0003284814883954823
step: 210, loss: 0.0003127541858702898
step: 220, loss: 7.783529144944623e-05
step: 230, loss: 0.00038276996929198503
step: 240, loss: 9.365245932713151e-05
step: 250, loss: 0.001259015523828566
step: 260, loss: 0.00015692257147748023
step: 270, loss: 0.07226871699094772
step: 280, loss: 0.06425037235021591
step: 290, loss: 0.0015931092202663422
step: 300, loss: 0.0020855963230133057
step: 310, loss: 0.004794291220605373
step: 320, loss: 0.017461709678173065
step: 330, loss: 0.004231874365359545
step: 340, loss: 0.01350106205791235
step: 350, loss: 0.000209203120903112
step: 360, loss: 0.0007482976652681828
step: 370, loss: 0.00044116380740888417
step: 380, loss: 0.006304113194346428
step: 390, loss: 0.00045896461233496666
step: 400, loss: 0.0005511585623025894
step: 410, loss: 0.007429542951285839
step: 420, loss: 0.0023654478136450052
step: 430, loss: 0.0002146952028851956
step: 440, loss: 0.015528281219303608
step: 450, loss: 0.022986235097050667
step: 460, loss: 0.0003616886679083109
step: 470, loss: 0.001961720874533057
step: 480, loss: 0.037142314016819
step: 490, loss: 0.004637132398784161

```

step: 500, loss: 0.000760572322178632
step: 510, loss: 0.0009669058490544558
step: 520, loss: 0.002510347403585911
step: 530, loss: 0.0006889255600981414
/usr/local/lib/python3.7/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=0.9344569288389513, f1=0.9273743016759776,
best_f1=0.9272137227630968
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040390403592027724
step: 10, loss: 0.0012508867075666785
step: 20, loss: 0.0003066679055336863
step: 30, loss: 0.0008388845017179847
step: 40, loss: 0.00034899963065981865
step: 50, loss: 0.012862395495176315
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.00043983079376630485
step: 70, loss: 0.0019150185398757458
step: 80, loss: 0.005262686405330896
step: 90, loss: 0.0015686313854530454
step: 100, loss: 0.00012585727381519973
step: 110, loss: 0.0002930688497144729
step: 120, loss: 0.00022717878164257854
step: 130, loss: 0.04022463038563728
step: 140, loss: 0.0006436927942559123
step: 150, loss: 0.007982085458934307
step: 160, loss: 0.0007033630390651524
step: 170, loss: 0.011321288533508778
step: 180, loss: 0.00018276067567057908
step: 190, loss: 0.003436551196500659
step: 200, loss: 0.0010426456574350595
step: 210, loss: 0.0016661732224747539
step: 220, loss: 0.002261901507154107
step: 230, loss: 5.906377555220388e-05
step: 240, loss: 7.716403342783451e-05
step: 250, loss: 0.00018538469157647341
step: 260, loss: 0.00015389721374958754
step: 270, loss: 0.0007102079107426107
step: 280, loss: 0.15041133761405945
step: 290, loss: 8.635130507173017e-05
step: 300, loss: 0.0007915357127785683
step: 310, loss: 0.0008638535509817302
step: 320, loss: 0.005249197129160166
step: 330, loss: 0.00026694100233726203
step: 340, loss: 8.421535312663764e-05

```

```

step: 350, loss: 8.863227412803099e-05
step: 360, loss: 4.762049138662405e-05
step: 370, loss: 0.009148281998932362
step: 380, loss: 0.0008901027613319457
step: 390, loss: 0.1392165571451187
step: 400, loss: 0.0037792143411934376
step: 410, loss: 0.00013547635171562433
step: 420, loss: 0.0009475798578932881
step: 430, loss: 0.006875597406178713
step: 440, loss: 0.010061689652502537
step: 450, loss: 0.0715487077832222
step: 460, loss: 0.002626709872856736
step: 470, loss: 0.003130408236756921
step: 480, loss: 0.000501624250318855
step: 490, loss: 0.007215987890958786
step: 500, loss: 0.00040152863948605955
step: 510, loss: 0.0002896194637287408
step: 520, loss: 0.008976859971880913
step: 530, loss: 4.169771636952646e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 8: dev_f1=0.9365760610395804, f1=0.9182990922121358,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011186351766809821
step: 10, loss: 5.2770723414141685e-05
step: 20, loss: 0.00017828747513704002
step: 30, loss: 0.0023996697273105383
step: 40, loss: 6.215899338712916e-05
step: 50, loss: 0.00014181082951836288
step: 60, loss: 5.381164373829961e-05
step: 70, loss: 0.0072986832819879055
step: 80, loss: 0.0004186795558780432
step: 90, loss: 9.923920879373327e-05
step: 100, loss: 0.00012999399041291326
step: 110, loss: 9.73199712461792e-05
step: 120, loss: 0.00036376024945639074
step: 130, loss: 0.03690587356686592
step: 140, loss: 0.0008125274325720966
step: 150, loss: 0.0024481110740453005
step: 160, loss: 0.0024127590004354715

```

```

step: 170, loss: 0.008649058640003204
step: 180, loss: 7.39467068342492e-05
step: 190, loss: 7.719209679635242e-05
step: 200, loss: 0.007457563187927008
step: 210, loss: 0.00010988880967488512
step: 220, loss: 0.00015618535690009594
step: 230, loss: 7.355539128184319e-05
step: 240, loss: 0.02644377388060093
step: 250, loss: 0.002143742283806205
step: 260, loss: 0.010439714416861534
step: 270, loss: 5.131130092195235e-05
step: 280, loss: 0.0013499314663931727
step: 290, loss: 0.005442557856440544
step: 300, loss: 0.00020636389672290534
step: 310, loss: 0.012571691535413265
step: 320, loss: 0.0008294770959764719
step: 330, loss: 0.0004383068298920989
step: 340, loss: 0.00023078659432940185
step: 350, loss: 0.00014859951625112444
step: 360, loss: 4.52155145467259e-05
step: 370, loss: 0.00020609673811122775
step: 380, loss: 0.0005565829342231154
step: 390, loss: 0.0001414047583239153
step: 400, loss: 0.002213717671111226
step: 410, loss: 0.00012286726268939674
step: 420, loss: 0.00011076685041189194
step: 430, loss: 0.0007056518807075918
step: 440, loss: 0.00033520793658681214
step: 450, loss: 0.00013420585310086608
step: 460, loss: 0.00010155110794585198
step: 470, loss: 6.63664031890221e-05
step: 480, loss: 5.751186836278066e-05
step: 490, loss: 4.837114829570055e-05
step: 500, loss: 0.001280138036236167
step: 510, loss: 0.0001685680908849463
step: 520, loss: 5.622846947517246e-05
step: 530, loss: 3.723657573573291e-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=0.9333954354913834, f1=0.9190697674418604,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017394457245245576
step: 10, loss: 4.08478663302958e-05
step: 20, loss: 2.9246441044961102e-05

```

step: 30, loss: 4.510571307037026e-05
step: 40, loss: 0.004033718258142471
step: 50, loss: 0.0002250542165711522
step: 60, loss: 4.76178975077346e-05
step: 70, loss: 5.504829823621549e-05
step: 80, loss: 6.963252963032573e-05
step: 90, loss: 7.713835657341406e-05
step: 100, loss: 3.12656193273142e-05
step: 110, loss: 0.00026002994854934514
step: 120, loss: 5.08709235873539e-05
step: 130, loss: 0.0001313033135375008
step: 140, loss: 0.00014214224938768893
step: 150, loss: 0.0005620788433589041
step: 160, loss: 8.671922114444897e-05
step: 170, loss: 0.00040948003879748285
step: 180, loss: 6.718875374644995e-05
step: 190, loss: 6.0241163737373427e-05
step: 200, loss: 0.00031455254065804183
step: 210, loss: 0.00015773935592733324
step: 220, loss: 0.0007074982859194279
step: 230, loss: 5.62215136596933e-05
step: 240, loss: 7.278079283423722e-05
step: 250, loss: 3.849815402645618e-05
step: 260, loss: 0.0005067503661848605
step: 270, loss: 0.0003540661418810487
step: 280, loss: 0.0004428609390743077
step: 290, loss: 5.4140058637131006e-05
step: 300, loss: 4.237448229105212e-05
step: 310, loss: 2.9201879442553036e-05
step: 320, loss: 0.006374639458954334
step: 330, loss: 2.4422441128990613e-05
step: 340, loss: 0.003272698726505041
step: 350, loss: 0.00017164355085697025
step: 360, loss: 0.00023345084628090262
step: 370, loss: 0.000378690951038152
step: 380, loss: 8.865267591318116e-05
step: 390, loss: 0.013853328302502632
step: 400, loss: 0.0002288996911374852
step: 410, loss: 0.0062864674255251884
step: 420, loss: 0.000358732882887125
step: 430, loss: 6.661711813649163e-05
step: 440, loss: 0.0009681205847300589
step: 450, loss: 4.37651397078298e-05
step: 460, loss: 7.874359289417043e-05
step: 470, loss: 9.780227992450818e-05
step: 480, loss: 5.1004113629460335e-05
step: 490, loss: 8.027883450267836e-05
step: 500, loss: 0.0002152318338630721


```

step: 510, loss: 5.5209253332577646e-05
step: 520, loss: 0.00017012089665513486
step: 530, loss: 0.0012972523691132665
/usr/local/lib/python3.7/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=0.9340866290018832, f1=0.9222065063649223,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027342357207089663
step: 10, loss: 5.3870127885602415e-05
step: 20, loss: 4.317060665925965e-05
step: 30, loss: 4.412170164869167e-05
step: 40, loss: 7.977334462339059e-05
step: 50, loss: 8.056686783675104e-05
step: 60, loss: 0.0012600923655554652
step: 70, loss: 0.00019471465202514082
step: 80, loss: 4.542526221484877e-05
step: 90, loss: 0.0001430870033800602
step: 100, loss: 0.0005008465377613902
step: 110, loss: 7.836832810426131e-05
step: 120, loss: 0.0002217122819274664
step: 130, loss: 0.00011003485269611701
step: 140, loss: 7.543046376667917e-05
step: 150, loss: 4.332789467298426e-05
step: 160, loss: 2.9231345251901075e-05
step: 170, loss: 2.9052936952211894e-05
step: 180, loss: 0.005519400350749493
step: 190, loss: 0.004197920206934214
step: 200, loss: 0.0014616854023188353
step: 210, loss: 0.0002201780880568549
step: 220, loss: 0.001126280054450035
step: 230, loss: 5.869208325748332e-05
step: 240, loss: 0.00028988183476030827
step: 250, loss: 2.2254469513427466e-05
step: 260, loss: 4.587876901496202e-05
step: 270, loss: 6.746464350726455e-05
step: 280, loss: 4.105461266590282e-05
step: 290, loss: 0.0004727281047962606
step: 300, loss: 7.211006595753133e-05
step: 310, loss: 3.818251934717409e-05
step: 320, loss: 7.657417882001027e-05
step: 330, loss: 5.571676592808217e-05
step: 340, loss: 0.005880813580006361
step: 350, loss: 6.49440917186439e-05
step: 360, loss: 0.00011049892054870725

```

```

step: 370, loss: 0.0010352933313697577
step: 380, loss: 6.191047577885911e-05
step: 390, loss: 8.686474757269025e-05
step: 400, loss: 5.803240492241457e-05
step: 410, loss: 3.8302114262478426e-05
step: 420, loss: 6.446677434723824e-05
step: 430, loss: 0.0001432519347872585
step: 440, loss: 0.00025766450562514365
step: 450, loss: 7.09254527464509e-05
step: 460, loss: 0.030401429161429405
step: 470, loss: 3.125400689896196e-05
step: 480, loss: 0.00011766709212679416
step: 490, loss: 0.0005571133806370199
step: 500, loss: 0.00043077568989247084
step: 510, loss: 8.485165744787082e-05
step: 520, loss: 5.752086144639179e-05
step: 530, loss: 0.010567037388682365
/usr/local/lib/python3.7/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=0.9311320754716981, f1=0.913556920170052,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.1879753805696964e-05
step: 10, loss: 2.0403138478286564e-05
step: 20, loss: 0.00041309266816824675
step: 30, loss: 0.0013160220114514232
step: 40, loss: 7.709916826570407e-05
step: 50, loss: 5.0484224630054086e-05
step: 60, loss: 0.00029329099925234914
step: 70, loss: 0.0002452540211379528
step: 80, loss: 0.002330988412722945
step: 90, loss: 0.00014332152204588056
step: 100, loss: 6.334352656267583e-05
step: 110, loss: 0.00018050607468467206
step: 120, loss: 2.3580225388286635e-05
step: 130, loss: 4.5169963414082304e-05
step: 140, loss: 3.7663881812477484e-05
step: 150, loss: 7.265512977028266e-05
step: 160, loss: 2.4701899747014977e-05
step: 170, loss: 2.9697246645810083e-05
step: 180, loss: 3.4625256375875324e-05
step: 190, loss: 4.7287863708334044e-05
step: 200, loss: 6.000322900945321e-05
step: 210, loss: 0.00016706521273590624
step: 220, loss: 3.281503086327575e-05

```

```

step: 230, loss: 2.0637806301238015e-05
step: 240, loss: 0.00018433821969665587
step: 250, loss: 2.2403424736694433e-05
step: 260, loss: 3.104880670434795e-05
step: 270, loss: 0.0033879110123962164
step: 280, loss: 0.00026323599740862846
step: 290, loss: 6.518009467981756e-05
step: 300, loss: 7.866616942919791e-05
step: 310, loss: 3.44087638950441e-05
step: 320, loss: 1.9006149159395136e-05
step: 330, loss: 0.00020936888176947832
step: 340, loss: 2.937238423328381e-05
step: 350, loss: 5.925962977926247e-05
step: 360, loss: 0.00012058016000082716
step: 370, loss: 3.403632581466809e-05
step: 380, loss: 2.564808528404683e-05
step: 390, loss: 0.06684788316488266
step: 400, loss: 5.1820639782818034e-05
step: 410, loss: 2.450409738230519e-05
step: 420, loss: 3.8755715650040656e-05
step: 430, loss: 3.336540248710662e-05
step: 440, loss: 3.286932042101398e-05
step: 450, loss: 3.126497540506534e-05
step: 460, loss: 1.8022719814325683e-05
step: 470, loss: 2.35354746109806e-05
step: 480, loss: 7.581948739243671e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 490, loss: 0.00015606862143613398
step: 500, loss: 0.00014785297389607877
step: 510, loss: 0.004302083980292082
step: 520, loss: 2.629238042572979e-05
step: 530, loss: 0.0527227520942688
/usr/local/lib/python3.7/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=0.9332706766917293, f1=0.9194756554307116,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019448879174888134
step: 10, loss: 5.273127317195758e-05
step: 20, loss: 5.725146547774784e-05
step: 30, loss: 2.9712047762586735e-05
step: 40, loss: 3.659473077277653e-05
step: 50, loss: 0.00011259539314778522
step: 60, loss: 4.321780215832405e-05
step: 70, loss: 0.00012042629532516003

```

```
step: 80, loss: 2.8195829145261087e-05
step: 90, loss: 0.0002210403181379661
step: 100, loss: 3.051640851481352e-05
step: 110, loss: 0.00018499558791518211
step: 120, loss: 5.44606227776967e-05
step: 130, loss: 2.5875122446450405e-05
step: 140, loss: 0.0017187207704409957
step: 150, loss: 6.000500434311107e-05
step: 160, loss: 3.8930043956497684e-05
step: 170, loss: 0.00011927523883059621
step: 180, loss: 5.129589408170432e-05
step: 190, loss: 0.0006644694949500263
step: 200, loss: 0.00031220263917930424
step: 210, loss: 2.9436341719701886e-05
step: 220, loss: 2.3487429643864743e-05
step: 230, loss: 8.778498886385933e-05
step: 240, loss: 0.025366811081767082
step: 250, loss: 6.611204298678786e-05
step: 260, loss: 3.353271313244477e-05
step: 270, loss: 3.079898306168616e-05
step: 280, loss: 3.79257726308424e-05
step: 290, loss: 6.945554923731834e-05
step: 300, loss: 8.277945016743615e-05
step: 310, loss: 0.008196452632546425
step: 320, loss: 0.0026862272061407566
step: 330, loss: 0.001310506951995194
step: 340, loss: 0.000143088647746481
step: 350, loss: 6.381948333000764e-05
step: 360, loss: 6.739344098605216e-05
step: 370, loss: 0.00020318813039921224
step: 380, loss: 0.0010298279812559485
step: 390, loss: 6.380579725373536e-05
step: 400, loss: 0.00021134479902684689
step: 410, loss: 0.00011270544928265736
step: 420, loss: 4.3529966205824167e-05
step: 430, loss: 6.686478445772082e-05
step: 440, loss: 7.803340122336522e-05
step: 450, loss: 0.00013601151295006275
step: 460, loss: 0.0007039060583338141
step: 470, loss: 4.173777051619254e-05
step: 480, loss: 2.7104437322122976e-05
step: 490, loss: 0.00016055155720096081
step: 500, loss: 1.786620487109758e-05
step: 510, loss: 6.885619222884998e-05
step: 520, loss: 0.0004826945951208472
step: 530, loss: 3.0549243092536926e-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=0.9319271332694151, f1=0.9091787439613526,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.234426786657423e-05
step: 10, loss: 2.002284054469783e-05
step: 20, loss: 2.7722559025278315e-05
step: 30, loss: 0.0001440558407921344
step: 40, loss: 5.1083708967780694e-05
step: 50, loss: 0.00024026080791372806
step: 60, loss: 2.752127511485014e-05
step: 70, loss: 4.0378225094173104e-05
step: 80, loss: 5.906034493818879e-05
step: 90, loss: 7.763072790112346e-05
step: 100, loss: 0.00014803174417465925
step: 110, loss: 7.033316796878353e-05
step: 120, loss: 8.73956159921363e-05
step: 130, loss: 0.00016287424659822136
step: 140, loss: 2.7126432541990653e-05
step: 150, loss: 8.625135524198413e-05
step: 160, loss: 4.191449988866225e-05
step: 170, loss: 4.357858415460214e-05
step: 180, loss: 2.819169094436802e-05
step: 190, loss: 0.00010242173448204994
step: 200, loss: 0.00017756350280251354
step: 210, loss: 2.8702308554784395e-05
step: 220, loss: 2.400515768385958e-05
step: 230, loss: 3.5533848858904094e-05
step: 240, loss: 2.9629623895743862e-05
step: 250, loss: 0.00038724884507246315
step: 260, loss: 2.9760236429865472e-05
step: 270, loss: 2.559942367952317e-05
step: 280, loss: 2.0034256522194482e-05
step: 290, loss: 0.00013978421338833869
step: 300, loss: 4.3723444832721725e-05
step: 310, loss: 0.0006004479364491999
step: 320, loss: 0.00015254398749675602
step: 330, loss: 6.364625005517155e-05
step: 340, loss: 3.342596028232947e-05
step: 350, loss: 7.049906707834452e-05
step: 360, loss: 0.00014691676187794656
step: 370, loss: 0.00043972552521154284
step: 380, loss: 4.381880353321321e-05
step: 390, loss: 0.03467344492673874
step: 400, loss: 0.0008444760460406542
step: 410, loss: 0.00025138119235634804
```

```

step: 420, loss: 3.338713213452138e-05
step: 430, loss: 0.00045610294910147786
step: 440, loss: 7.940283830976114e-05
step: 450, loss: 0.00011491176701383665
step: 460, loss: 8.377804624615237e-05
step: 470, loss: 5.7138080592267215e-05
step: 480, loss: 3.8744859921280295e-05
step: 490, loss: 4.3352665670681745e-05
step: 500, loss: 3.8322930777212605e-05
step: 510, loss: 0.0006206388934515417
step: 520, loss: 3.2441657822346315e-05
step: 530, loss: 8.565666939830408e-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=0.9338446788111218, f1=0.9165067178502879,
best_f1=0.9182990922121358
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7391162802814506e-05
step: 10, loss: 3.8738005969207734e-05
step: 20, loss: 0.00012495825649239123
step: 30, loss: 0.00016951403813436627
step: 40, loss: 0.06681101024150848
step: 50, loss: 0.00010557743371464312
step: 60, loss: 1.579874697199557e-05
step: 70, loss: 7.924111559987068e-05
step: 80, loss: 7.739248394500464e-05
step: 90, loss: 3.347011806908995e-05
step: 100, loss: 8.664101187605411e-05
step: 110, loss: 0.0005122966831550002
step: 120, loss: 0.00041538619552738965
step: 130, loss: 0.0070480480790138245
step: 140, loss: 2.8437512810342014e-05
step: 150, loss: 0.00023115449585020542
step: 160, loss: 2.3230277292896062e-05
step: 170, loss: 4.891716525889933e-05
step: 180, loss: 1.87042878678767e-05
step: 190, loss: 5.7657231081975624e-05
step: 200, loss: 3.224020838388242e-05
step: 210, loss: 0.000647236593067646
step: 220, loss: 2.0037874492118135e-05
step: 230, loss: 0.0011330011766403913
step: 240, loss: 0.00019822495232801884
step: 250, loss: 3.234888208680786e-05
step: 260, loss: 7.335448754020035e-05
step: 270, loss: 2.9755989089608192e-05

```

```

step: 280, loss: 0.0003443742170929909
step: 290, loss: 7.206398731796071e-05
step: 300, loss: 2.1475734683917835e-05
step: 310, loss: 3.1574083550367504e-05
step: 320, loss: 3.829310662695207e-05
step: 330, loss: 0.0011340021155774593
step: 340, loss: 1.8216302123619244e-05
step: 350, loss: 2.4528850190108642e-05
step: 360, loss: 0.0005182178574614227
step: 370, loss: 9.943562326952815e-05
step: 380, loss: 0.0002598853316158056
step: 390, loss: 0.0008547154720872641
step: 400, loss: 3.549905159161426e-05
step: 410, loss: 2.406461317150388e-05
step: 420, loss: 3.0151382816256955e-05
step: 430, loss: 0.022699952125549316
step: 440, loss: 5.0309965445194393e-05
step: 450, loss: 0.00012688766582868993
step: 460, loss: 3.2382289646193385e-05
step: 470, loss: 0.002060920000076294
step: 480, loss: 2.1330033632693812e-05
step: 490, loss: 2.4757573555689305e-05
step: 500, loss: 3.428530544624664e-05
step: 510, loss: 6.172592838993296e-05
step: 520, loss: 2.1457293769344687e-05
step: 530, loss: 2.8538643164210953e-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=0.9355608591885441, f1=0.9190225203641591,
best_f1=0.9182990922121358
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 [00:24, 235.40it/s]
load_f1 = 0.9366130558183537
real_f1 = 0.934916864608076
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 [00:18, 235.28it/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 bert-base-uncased \
--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 bert-base-uncased \
--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: 0.563360333442688
```

```
step: 10, loss: 0.3503831624984741
```

```

step: 20, loss: 0.4105105698108673
step: 30, loss: 0.3261139690876007
step: 40, loss: 0.20120301842689514
step: 50, loss: 0.4789356291294098
step: 60, loss: 0.34300094842910767
step: 70, loss: 0.2249271273612976
step: 80, loss: 0.18367575109004974
step: 90, loss: 0.400393545627594
step: 100, loss: 0.41239821910858154
step: 110, loss: 0.2505453824996948
step: 120, loss: 0.2032642960548401
step: 130, loss: 0.2150827944278717
step: 140, loss: 0.22916072607040405
step: 150, loss: 0.22421923279762268
step: 160, loss: 0.23917867243289948
step: 170, loss: 0.31388622522354126
step: 180, loss: 0.117743581533432
step: 190, loss: 0.2402559220790863
step: 200, loss: 0.2509887218475342
step: 210, loss: 0.25570595264434814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.540268456375839, f1=0.5800000000000001,
best_f1=0.5800000000000001
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17664921283721924
step: 10, loss: 0.27837231755256653
step: 20, loss: 0.22109819948673248
step: 30, loss: 0.25786665081977844
step: 40, loss: 0.2745864987373352
step: 50, loss: 0.21357087790966034
step: 60, loss: 0.3675594925880432
step: 70, loss: 0.11966989189386368
step: 80, loss: 0.2251892387866974
step: 90, loss: 0.10733044892549515
step: 100, loss: 0.013442235998809338
step: 110, loss: 0.11826404184103012
step: 120, loss: 0.11911533027887344
step: 130, loss: 0.04031048342585564
step: 140, loss: 0.2616669237613678
step: 150, loss: 0.193214550614357

```

```

step: 160, loss: 0.194614976644516
step: 170, loss: 0.11580502986907959
step: 180, loss: 0.1965094953775406
step: 190, loss: 0.22741568088531494
step: 200, loss: 0.04976562783122063
step: 210, loss: 0.12227512151002884
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5963636363636364, f1=0.6329588014981273,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03005756437778473
step: 10, loss: 0.21435211598873138
step: 20, loss: 0.22298549115657806
step: 30, loss: 0.05987906455993652
step: 40, loss: 0.14741632342338562
step: 50, loss: 0.1451779454946518
step: 60, loss: 0.05692560598254204
step: 70, loss: 0.08794037252664566
step: 80, loss: 0.17160925269126892
step: 90, loss: 0.03592577949166298
step: 100, loss: 0.11533746123313904
step: 110, loss: 0.1730012744665146
step: 120, loss: 0.1340017169713974
step: 130, loss: 0.17459821701049805
step: 140, loss: 0.1265365481376648
step: 150, loss: 0.15194495022296906
step: 160, loss: 0.013118579052388668
step: 170, loss: 0.1584625244140625
step: 180, loss: 0.049523405730724335
step: 190, loss: 0.20823487639427185
step: 200, loss: 0.04815278947353363
step: 210, loss: 0.1967179775238037
/usr/local/lib/python3.7/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=0.5759162303664921, f1=0.6143344709897611,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.08072014898061752
step: 10, loss: 0.07716238498687744
step: 20, loss: 0.03119313344359398
step: 30, loss: 0.08885606378316879
step: 40, loss: 0.042767200618982315
step: 50, loss: 0.10173439234495163
step: 60, loss: 0.10029321908950806
step: 70, loss: 0.23678088188171387
step: 80, loss: 0.127373605966568
step: 90, loss: 0.015954140573740005
step: 100, loss: 0.1958421915769577
step: 110, loss: 0.07008136063814163
step: 120, loss: 0.07413958758115768
step: 130, loss: 0.09217330068349838
step: 140, loss: 0.21308274567127228
step: 150, loss: 0.03564954921603203
step: 160, loss: 0.02336619235575199
step: 170, loss: 0.0968162938952446
step: 180, loss: 0.1843322366476059
step: 190, loss: 0.09285930544137955
step: 200, loss: 0.15167710185050964
step: 210, loss: 0.1945231705904007
/usr/local/lib/python3.7/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=0.5870841487279843, f1=0.6363636363636365,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11068297177553177
step: 10, loss: 0.018535101786255836
step: 20, loss: 0.04281897097826004
step: 30, loss: 0.022144030779600143
step: 40, loss: 0.025472577661275864
step: 50, loss: 0.04114553704857826
step: 60, loss: 0.054316021502017975
step: 70, loss: 0.06000697240233421
step: 80, loss: 0.1430298238992691
step: 90, loss: 0.16454939544200897
step: 100, loss: 0.002541961148381233
step: 110, loss: 0.10301757603883743
step: 120, loss: 0.058145541697740555
step: 130, loss: 0.07875803858041763
step: 140, loss: 0.10440153628587723
step: 150, loss: 0.050497423857450485
step: 160, loss: 0.1507968306541443
step: 170, loss: 0.04902178794145584

```

```

step: 180, loss: 0.053887102752923965
step: 190, loss: 0.006940164603292942
step: 200, loss: 0.15851211547851562
step: 210, loss: 0.007250573951750994
/usr/local/lib/python3.7/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=0.5585215605749486, f1=0.6147540983606558,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020463505759835243
step: 10, loss: 0.05030268803238869
step: 20, loss: 0.0022997339256107807
step: 30, loss: 0.0006486869533546269
step: 40, loss: 0.09302273392677307
step: 50, loss: 0.0035654231905937195
step: 60, loss: 0.018637992441654205
step: 70, loss: 0.08970974385738373
step: 80, loss: 0.08950715512037277
step: 90, loss: 0.194692924618721
step: 100, loss: 0.032635707408189774
step: 110, loss: 0.021840829402208328
step: 120, loss: 0.026198435574769974
step: 130, loss: 0.0856182873249054
step: 140, loss: 0.022725556045770645
step: 150, loss: 0.016028767451643944
step: 160, loss: 0.006460533011704683
step: 170, loss: 0.07752082496881485
step: 180, loss: 0.012287558056414127
step: 190, loss: 0.04696967825293541
step: 200, loss: 0.004480244591832161
step: 210, loss: 0.045179061591625214
/usr/local/lib/python3.7/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=0.5641025641025642, f1=0.6057906458797327,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005138831678777933
step: 10, loss: 0.008256569504737854
step: 20, loss: 0.011069356463849545
step: 30, loss: 0.037517037242650986
step: 40, loss: 0.0317300409078598
step: 50, loss: 0.04713372141122818

```

```

step: 60, loss: 0.009419750422239304
step: 70, loss: 0.012600899673998356
step: 80, loss: 0.02642858773469925
step: 90, loss: 0.007969693280756474
step: 100, loss: 0.003431286895647645
step: 110, loss: 0.09477470815181732
step: 120, loss: 0.06173530966043472
step: 130, loss: 0.017004158347845078
step: 140, loss: 0.0296366848051548
step: 150, loss: 0.0418255478143692
step: 160, loss: 0.013520305044949055
step: 170, loss: 0.0235359538346529
step: 180, loss: 0.026575801894068718
step: 190, loss: 0.10815361887216568
step: 200, loss: 0.05452549457550049
step: 210, loss: 0.003038847353309393
/usr/local/lib/python3.7/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=0.5763440860215053, f1=0.6313645621181263,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00435477402061224
step: 10, loss: 0.0425267219543457
step: 20, loss: 0.002241868991404772
step: 30, loss: 0.03269007056951523
step: 40, loss: 0.019471708685159683
step: 50, loss: 0.0021740440279245377
step: 60, loss: 0.11257948726415634
step: 70, loss: 0.021384116262197495
step: 80, loss: 0.015801405534148216
step: 90, loss: 0.09447604417800903
step: 100, loss: 0.006964368280023336
step: 110, loss: 0.04490488022565842
step: 120, loss: 0.0010223345598205924
step: 130, loss: 0.003523328574374318
step: 140, loss: 0.09991740435361862
step: 150, loss: 0.031559739261865616
step: 160, loss: 0.008494120091199875
step: 170, loss: 0.002063456689938903
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 180, loss: 0.12886695563793182
step: 190, loss: 0.17302480340003967
step: 200, loss: 0.0025183469988405704
step: 210, loss: 0.04165862500667572
/usr/local/lib/python3.7/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=0.579520697167756, f1=0.6322580645161291,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011400529183447361
step: 10, loss: 0.004235875327140093
step: 20, loss: 0.0007694929954595864
step: 30, loss: 0.001286288839764893
step: 40, loss: 0.05214017257094383
step: 50, loss: 0.0005088350153528154
step: 60, loss: 0.013371741399168968
step: 70, loss: 0.0003622269432526082
step: 80, loss: 0.0010336491977795959
step: 90, loss: 0.0012933410471305251
step: 100, loss: 0.023528529331088066
step: 110, loss: 0.03459262102842331
step: 120, loss: 0.0013804389163851738
step: 130, loss: 0.005207001231610775
step: 140, loss: 0.009093528613448143
step: 150, loss: 0.022521112114191055
step: 160, loss: 0.0006001211586408317
step: 170, loss: 0.0009101805626414716
step: 180, loss: 0.07498621195554733
step: 190, loss: 0.00026070926105603576
step: 200, loss: 0.0824221670627594
step: 210, loss: 0.004690409172326326
/usr/local/lib/python3.7/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=0.5720524017467249, f1=0.6315789473684211,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011501987464725971
step: 10, loss: 0.011343899182975292
step: 20, loss: 0.001848881016485393
step: 30, loss: 0.00961315631866455
step: 40, loss: 0.010896777734160423
step: 50, loss: 0.0010800468735396862
step: 60, loss: 0.038900237530469894
step: 70, loss: 0.005916444119066
step: 80, loss: 0.10831131041049957
step: 90, loss: 0.013077495619654655
step: 100, loss: 0.03601592034101486

```

```

step: 110, loss: 0.0002706335799302906
step: 120, loss: 0.0015133823035284877
step: 130, loss: 0.004909981973469257
step: 140, loss: 0.0006069550872780383
step: 150, loss: 0.003982912749052048
step: 160, loss: 0.0008774717571213841
step: 170, loss: 0.0008841177332215011
step: 180, loss: 0.018130086362361908
step: 190, loss: 0.17782141268253326
step: 200, loss: 0.007179387379437685
step: 210, loss: 0.09476633369922638
/usr/local/lib/python3.7/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=0.5783664459161149, f1=0.6411889596602973,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01110947597771883
step: 10, loss: 0.04825994372367859
step: 20, loss: 0.00614469638094306
step: 30, loss: 0.004864180460572243
step: 40, loss: 0.021005835384130478
step: 50, loss: 0.06448047608137131
step: 60, loss: 0.012686456553637981
step: 70, loss: 0.03266415372490883
step: 80, loss: 0.026104165241122246
step: 90, loss: 0.019069209694862366
step: 100, loss: 0.055845096707344055
step: 110, loss: 0.052059076726436615
step: 120, loss: 0.014618927612900734
step: 130, loss: 0.00560676259920001
step: 140, loss: 0.0005820785299874842
step: 150, loss: 0.00022644006821792573
step: 160, loss: 0.02425299771130085
step: 170, loss: 0.038971733301877975
step: 180, loss: 0.0007057894836179912
step: 190, loss: 0.0007205406436696649
step: 200, loss: 0.006359575316309929
step: 210, loss: 0.0010351981036365032
/usr/local/lib/python3.7/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=0.5663716814159292, f1=0.6458333333333333,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.053702957928180695
step: 10, loss: 0.00017244499758817255
step: 20, loss: 0.004252160899341106
step: 30, loss: 0.0025366938207298517
step: 40, loss: 0.005107346456497908
step: 50, loss: 0.0008793707238510251
step: 60, loss: 0.0014618179993703961
step: 70, loss: 0.0017720786854624748
step: 80, loss: 0.00042871941695921123
step: 90, loss: 0.003383324481546879
step: 100, loss: 0.007137601263821125
step: 110, loss: 0.00034173380117863417
step: 120, loss: 0.0023642953019589186
step: 130, loss: 6.57185009913519e-05
step: 140, loss: 0.00041210404015146196
step: 150, loss: 0.0006515298737213016
step: 160, loss: 0.004012248013168573
step: 170, loss: 0.030026167631149292
step: 180, loss: 0.002062223618850112
step: 190, loss: 0.014580950140953064
step: 200, loss: 0.0005191928939893842
step: 210, loss: 0.008644786663353443
/usr/local/lib/python3.7/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=0.5581395348837209, f1=0.609271523178808,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005174342077225447
step: 10, loss: 9.903189493343234e-05
step: 20, loss: 0.01143078226596117
step: 30, loss: 0.03384726494550705
step: 40, loss: 0.0007055692258290946
step: 50, loss: 0.0004784493939951062
step: 60, loss: 0.00019864653586409986
step: 70, loss: 0.0019857981242239475
step: 80, loss: 0.0010491838911548257
step: 90, loss: 0.0002400741068413481
step: 100, loss: 0.0013114864705130458
step: 110, loss: 0.00029726704815402627
step: 120, loss: 0.00018670885765459388
step: 130, loss: 0.014299076050519943
step: 140, loss: 0.004616493359208107
step: 150, loss: 0.017269032076001167
step: 160, loss: 0.03275209292769432

```

```

step: 170, loss: 0.0028466694056987762
step: 180, loss: 0.012071754783391953
step: 190, loss: 0.0003674958716146648
step: 200, loss: 0.0007314502145163715
step: 210, loss: 0.005525875836610794
/usr/local/lib/python3.7/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=0.5605381165919282, f1=0.6111111111111112,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.960497845895588e-05
step: 10, loss: 0.0650465115904808
step: 20, loss: 0.0001306365302298218
step: 30, loss: 0.032783206552267075
step: 40, loss: 0.00018337820074521005
step: 50, loss: 0.018666144460439682
step: 60, loss: 9.595046140020713e-05
step: 70, loss: 0.00010438133176648989
step: 80, loss: 0.00036616038414649665
step: 90, loss: 0.008816235698759556
step: 100, loss: 0.0031616908963769674
step: 110, loss: 0.0008806888945400715
step: 120, loss: 0.0002182106691179797
step: 130, loss: 0.0037130825221538544
step: 140, loss: 0.002939556958153844
step: 150, loss: 0.00035331628168933094
step: 160, loss: 0.00017482455587014556
step: 170, loss: 0.0001441782369511202
step: 180, loss: 0.0020997868850827217
step: 190, loss: 0.0002545146271586418
step: 200, loss: 0.00045864927233196795
step: 210, loss: 0.0006936974823474884
/usr/local/lib/python3.7/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=0.5540540540540541, f1=0.6226012793176973,
best_f1=0.6329588014981273
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001240027922904119
step: 10, loss: 0.0046764821745455265
step: 20, loss: 0.0007616077200509608
step: 30, loss: 0.005556656047701836
step: 40, loss: 0.0006691452581435442

```

```

step: 50, loss: 0.0002182352327508852
step: 60, loss: 0.000779944472014904
step: 70, loss: 0.0019263973226770759
step: 80, loss: 0.00033050504862330854
step: 90, loss: 0.021520795300602913
step: 100, loss: 0.0003443534660618752
step: 110, loss: 0.00017763575306162238
step: 120, loss: 0.00024243525695055723
step: 130, loss: 0.0030066706240177155
step: 140, loss: 0.00031715244404040277
step: 150, loss: 0.0002593204553704709
step: 160, loss: 0.012510450556874275
step: 170, loss: 0.0004771027306560427
step: 180, loss: 0.00018896773690357804
step: 190, loss: 0.00011549473856575787
step: 200, loss: 0.003694762010127306
step: 210, loss: 0.0008528839680366218
/usr/local/lib/python3.7/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=0.5585585585585586, f1=0.6268656716417911,
best_f1=0.6329588014981273
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 [00:05, 403.38it/s]
load_f1 = 0.582120582120582
real_f1 = 0.5793991416309012
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:19, 227.56it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5452269315719604
step: 10, loss: 0.36817097663879395
step: 20, loss: 0.28986942768096924
step: 30, loss: 0.45134422183036804
step: 40, loss: 0.4287320077419281
step: 50, loss: 0.30328235030174255
step: 60, loss: 0.2633335590362549
step: 70, loss: 0.2903966009616852
step: 80, loss: 0.23985759913921356
step: 90, loss: 0.31227874755859375
step: 100, loss: 0.32209303975105286
step: 110, loss: 0.37393274903297424
```

```

step: 120, loss: 0.2303493469953537
step: 130, loss: 0.17246727645397186
step: 140, loss: 0.04297035560011864
step: 150, loss: 0.17568805813789368
step: 160, loss: 0.15779371559619904
step: 170, loss: 0.17679567635059357
step: 180, loss: 0.04662597179412842
step: 190, loss: 0.21532569825649261
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6683804627249357, f1=0.645, best_f1=0.645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2817363142967224
step: 10, loss: 0.05807296931743622
step: 20, loss: 0.2248144894838333
step: 30, loss: 0.14834201335906982
step: 40, loss: 0.22870375216007233
step: 50, loss: 0.08230677247047424
step: 60, loss: 0.2019953578710556
step: 70, loss: 0.1251639574766159
step: 80, loss: 0.2000456601381302
step: 90, loss: 0.33416107296943665
step: 100, loss: 0.026996886357665062
step: 110, loss: 0.32853156328201294
step: 120, loss: 0.21320727467536926
step: 130, loss: 0.10905258357524872
step: 140, loss: 0.044592246413230896
step: 150, loss: 0.10344403982162476
step: 160, loss: 0.0951179563999176
step: 170, loss: 0.14392924308776855
step: 180, loss: 0.2426450103521347
step: 190, loss: 0.06577770411968231
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7396449704142012, f1=0.7457627118644068,
best_f1=0.7457627118644068

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10509008169174194
step: 10, loss: 0.040476422756910324
step: 20, loss: 0.13620203733444214
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.04464838281273842
step: 40, loss: 0.036566730588674545
step: 50, loss: 0.1962156891822815
step: 60, loss: 0.056676074862480164
step: 70, loss: 0.11589344590902328
step: 80, loss: 0.19252638518810272
step: 90, loss: 0.10209353268146515
step: 100, loss: 0.11452794075012207
step: 110, loss: 0.05633218213915825
step: 120, loss: 0.023028751835227013
step: 130, loss: 0.011186822317540646
step: 140, loss: 0.00762784481048584
step: 150, loss: 0.11540187895298004
step: 160, loss: 0.013530520722270012
step: 170, loss: 0.15212246775627136
step: 180, loss: 0.022288713604211807
step: 190, loss: 0.12219218164682388
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7671957671957672, f1=0.7692307692307692,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012715582735836506
step: 10, loss: 0.06311271339654922
step: 20, loss: 0.09582587331533432
step: 30, loss: 0.03351597487926483
step: 40, loss: 0.02944793738424778
step: 50, loss: 0.0509248748421669
step: 60, loss: 0.11829254031181335
step: 70, loss: 0.047074396163225174
step: 80, loss: 0.05815267562866211
step: 90, loss: 0.01310658548027277
step: 100, loss: 0.11931125819683075
step: 110, loss: 0.0034231669269502163

```

```

step: 120, loss: 0.028076669201254845
step: 130, loss: 0.24677535891532898
step: 140, loss: 0.010788812302052975
step: 150, loss: 0.023105179890990257
step: 160, loss: 0.017650820314884186
step: 170, loss: 0.039781734347343445
step: 180, loss: 0.0015132622793316841
step: 190, loss: 0.13045012950897217
/usr/local/lib/python3.7/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=0.7300771208226221, f1=0.7407407407407407,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07308498024940491
step: 10, loss: 0.0033747439738363028
step: 20, loss: 0.007382153999060392
step: 30, loss: 0.004429445136338472
step: 40, loss: 0.092521533370018
step: 50, loss: 0.023120488971471786
step: 60, loss: 0.15245865285396576
step: 70, loss: 0.047948792576789856
step: 80, loss: 0.017886372283101082
step: 90, loss: 0.02656984142959118
step: 100, loss: 0.0007936947513371706
step: 110, loss: 0.0036146086640655994
step: 120, loss: 0.001668113050982356
step: 130, loss: 0.00793603714555502
step: 140, loss: 0.07555221021175385
step: 150, loss: 0.00561038963496685
step: 160, loss: 0.048836614936590195
step: 170, loss: 0.039414625614881516
step: 180, loss: 0.054794829338788986
step: 190, loss: 0.07125131785869598
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 5: dev_f1=0.7310704960835509, f1=0.7262872628726287,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0031970273703336716
step: 10, loss: 0.020742442458868027
step: 20, loss: 0.006146089173853397
step: 30, loss: 0.00038968416629359126

```



```

step: 40, loss: 0.004602793604135513
step: 50, loss: 0.002848108299076557
step: 60, loss: 0.06474051624536514
step: 70, loss: 0.0784892737865448
step: 80, loss: 0.032977793365716934
step: 90, loss: 0.005671176128089428
step: 100, loss: 0.001222697552293539
step: 110, loss: 0.002505709184333682
step: 120, loss: 0.04362298548221588
step: 130, loss: 0.016597097739577293
step: 140, loss: 0.09415131062269211
step: 150, loss: 0.007364785764366388
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 160, loss: 0.007079660426825285
step: 170, loss: 0.018650414422154427
step: 180, loss: 0.009884619154036045
step: 190, loss: 0.011729124002158642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.7311827956989247, f1=0.7374301675977653,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017322899075224996
step: 10, loss: 0.0006093577831052244
step: 20, loss: 0.0027315542101860046
step: 30, loss: 0.00561560969799757
step: 40, loss: 0.0010077308397740126
step: 50, loss: 0.0027419214602559805
step: 60, loss: 0.0013190245954319835
step: 70, loss: 0.0008036476792767644
step: 80, loss: 0.008132938295602798
step: 90, loss: 0.0007770649972371757
step: 100, loss: 0.00033500700374133885
step: 110, loss: 0.0020510314498096704
step: 120, loss: 0.0006311756442300975
step: 130, loss: 0.00035739943268708885
step: 140, loss: 0.000441909913206473
step: 150, loss: 0.036979738622903824
step: 160, loss: 0.18015483021736145
step: 170, loss: 0.003021097742021084
step: 180, loss: 0.0006977110751904547
step: 190, loss: 0.002377202734351158
/usr/local/lib/python3.7/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=0.7354497354497355, f1=0.7582417582417582,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026396203320473433
step: 10, loss: 0.01117573119699955
step: 20, loss: 0.005541789345443249
step: 30, loss: 0.005935350898653269
step: 40, loss: 0.0024840575642883778
step: 50, loss: 0.008471054956316948
step: 60, loss: 0.0024205902591347694
step: 70, loss: 0.008065887726843357
step: 80, loss: 0.0003527947992552072
step: 90, loss: 0.00035175096127204597
step: 100, loss: 0.03161201253533363
step: 110, loss: 0.000375796458683908
step: 120, loss: 0.009610197506844997
step: 130, loss: 0.0006875402759760618
step: 140, loss: 0.0024981095921248198
step: 150, loss: 0.0012344500282779336
step: 160, loss: 0.00041528273141011596
step: 170, loss: 0.0011643391335383058
step: 180, loss: 0.00978743378072977
step: 190, loss: 0.0005060415132902563
/usr/local/lib/python3.7/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=0.7346938775510203, f1=0.7636363636363637,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009534706478007138
step: 10, loss: 0.04930378124117851
step: 20, loss: 0.05107381194829941
step: 30, loss: 0.00020598860282916576
step: 40, loss: 0.0003552387352101505
step: 50, loss: 0.00028842836036346853
step: 60, loss: 0.0017900166567415
step: 70, loss: 0.0028451166581362486
step: 80, loss: 0.00039572990499436855
step: 90, loss: 0.0002675779105629772
step: 100, loss: 0.0019520105561241508
step: 110, loss: 0.03014741837978363
step: 120, loss: 0.0008956866804510355
step: 130, loss: 0.0217701718211174
step: 140, loss: 0.000519098190125078

```

```

step: 150, loss: 0.0035363254137337208
step: 160, loss: 0.0002697481832001358
step: 170, loss: 0.0003260981466155499
step: 180, loss: 0.0019500170601531863
step: 190, loss: 0.00016161489475052804
/usr/local/lib/python3.7/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=0.7540106951871658, f1=0.772117962466488,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023005357070360333
step: 10, loss: 0.00038503302494063973
step: 20, loss: 0.01875758171081543
step: 30, loss: 9.918970317812636e-05
step: 40, loss: 0.010479462333023548
step: 50, loss: 0.00028338286210782826
step: 60, loss: 0.0010597279760986567
step: 70, loss: 0.00023175703245215118
step: 80, loss: 0.00013418936578091234
step: 90, loss: 0.0004345369234215468
step: 100, loss: 0.0014630656223744154
step: 110, loss: 0.00038244135794229805
step: 120, loss: 0.01770181581377983
step: 130, loss: 0.008964535780251026
step: 140, loss: 0.007033176254481077
step: 150, loss: 0.1306622475385666
step: 160, loss: 0.0009637553594075143
step: 170, loss: 0.07725829631090164
step: 180, loss: 0.00035228158230893314
step: 190, loss: 0.0018447600305080414
/usr/local/lib/python3.7/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=0.765498652291105, f1=0.7828418230563003,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007687179022468626
step: 10, loss: 0.0005987334880046546
step: 20, loss: 0.04350616782903671
step: 30, loss: 0.06733270734548569
step: 40, loss: 0.0005106751923449337
step: 50, loss: 0.000783177965786308
step: 60, loss: 0.0007641083211638033

```

```

step: 70, loss: 0.00032885593827813864
step: 80, loss: 0.0004727505729533732
step: 90, loss: 0.0004547291318885982
step: 100, loss: 0.0007202949491329491
step: 110, loss: 0.00022641487885266542
step: 120, loss: 0.0004152871260885149
step: 130, loss: 0.002339035039767623
step: 140, loss: 0.0011024392442777753
step: 150, loss: 0.00029399959021247923
step: 160, loss: 0.0006650452269241214
step: 170, loss: 0.001004343619570136
step: 180, loss: 0.0007003378123044968
step: 190, loss: 0.00019151443848386407
/usr/local/lib/python3.7/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=0.7486910994764396, f1=0.768, best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010649684554664418
step: 10, loss: 0.01172295305877924
step: 20, loss: 0.00027816148940473795
step: 30, loss: 0.0007130308076739311
step: 40, loss: 0.00524082500487566
step: 50, loss: 0.0017109045293182135
step: 60, loss: 0.016006946563720703
step: 70, loss: 0.0001259191776625812
step: 80, loss: 0.000165314253536053
step: 90, loss: 0.00015187384269665927
step: 100, loss: 0.0001957739586941898
step: 110, loss: 0.00019068457186222076
step: 120, loss: 0.0002171052765334025
step: 130, loss: 0.00015991598775144666
step: 140, loss: 0.0005177899147383869
step: 150, loss: 0.00023508761660195887
step: 160, loss: 0.0001504052779637277
step: 170, loss: 0.002417949726805091
step: 180, loss: 0.00130832742433995
step: 190, loss: 0.00015640369383618236
/usr/local/lib/python3.7/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=0.736, f1=0.7603305785123966, best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00029612184152938426

```

```

step: 10, loss: 9.958762530004606e-05
step: 20, loss: 0.005426805000752211
step: 30, loss: 0.00017278778250329196
step: 40, loss: 0.00018404207366984338
step: 50, loss: 0.00044853572035208344
step: 60, loss: 0.00017731158004608005
step: 70, loss: 0.1745568960905075
step: 80, loss: 0.011667163111269474
step: 90, loss: 0.00027290545403957367
step: 100, loss: 0.00018345267744734883
step: 110, loss: 0.000208265075343661
step: 120, loss: 0.0006985389045439661
step: 130, loss: 0.0001649160694796592
step: 140, loss: 0.00013386792852543294
step: 150, loss: 9.396355744684115e-05
step: 160, loss: 0.0001951050799107179
step: 170, loss: 0.0006337066297419369
step: 180, loss: 0.0005257653538137674
step: 190, loss: 0.0011371728032827377
/usr/local/lib/python3.7/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=0.7602905569007263, f1=0.7493917274939174,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005261329002678394
step: 10, loss: 0.00025195899070240557
step: 20, loss: 0.00017727434169501066
step: 30, loss: 0.0045271990820765495
step: 40, loss: 0.005873749498277903
step: 50, loss: 0.00016686221351847053
step: 60, loss: 0.019923441112041473
step: 70, loss: 0.0001452265860280022
step: 80, loss: 0.00012880393478553742
step: 90, loss: 0.08874945342540741
step: 100, loss: 0.00028692357591353357
step: 110, loss: 0.0002707724052015692
step: 120, loss: 0.00016176192730199546
step: 130, loss: 0.00023900228552520275
step: 140, loss: 0.0011958708055317402
step: 150, loss: 0.00020032792235724628
step: 160, loss: 0.0027678448241204023
step: 170, loss: 0.00024268400738947093
step: 180, loss: 0.0007056510075926781
step: 190, loss: 0.0003931655955966562
/usr/local/lib/python3.7/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=0.7571801566579635, f1=0.7626666666666667,
best_f1=0.7692307692307692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007518379716202617
step: 10, loss: 0.00017904079868458211
step: 20, loss: 0.00017243711045011878
step: 30, loss: 0.0001055002212524414
step: 40, loss: 0.0002810796722769737
step: 50, loss: 0.00018935247499030083
step: 60, loss: 0.00016862331540323794
step: 70, loss: 0.0003353011270519346
step: 80, loss: 0.00022666265431325883
step: 90, loss: 0.00041081852396018803
step: 100, loss: 0.00038401797064580023
step: 110, loss: 0.0002597900456748903
step: 120, loss: 0.00013960196520201862
step: 130, loss: 0.00044668070040643215
step: 140, loss: 0.0003846963809337467
step: 150, loss: 0.0003018486895598471
step: 160, loss: 0.00015293850447051227
step: 170, loss: 0.0003429329372011125
step: 180, loss: 0.0019687237218022346
step: 190, loss: 0.0001735844707582146
/usr/local/lib/python3.7/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=0.751269035532995, f1=0.7552083333333333,
best_f1=0.7692307692307692
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale           : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:10, 204.71it/s]
load_f1 = 0.6253521126760564
real_f1 = 0.6022099447513812
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 [00:19, 224.55it/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 bert-base-uncased \
    --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 bert-base-uncased \
  --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: 0.632739782333374
step: 10, loss: 0.37038320302963257
step: 20, loss: 0.3050107955932617
step: 30, loss: 0.38223764300346375
step: 40, loss: 0.2891567051410675
step: 50, loss: 0.26228275895118713
step: 60, loss: 0.21775288879871368
step: 70, loss: 0.3533221483230591
step: 80, loss: 0.36923354864120483
step: 90, loss: 0.24865815043449402
step: 100, loss: 0.1936742216348648
step: 110, loss: 0.24791541695594788
step: 120, loss: 0.20225240290164948
step: 130, loss: 0.0640677809715271
step: 140, loss: 0.16259326040744781
step: 150, loss: 0.42403897643089294
step: 160, loss: 0.10192075371742249
step: 170, loss: 0.30268722772598267
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6860158311345647, f1=0.6701030927835051,
best_f1=0.6701030927835051
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18410290777683258
step: 10, loss: 0.1628885716199875
step: 20, loss: 0.19318018853664398
step: 30, loss: 0.2092280089855194
step: 40, loss: 0.0585927739739418
step: 50, loss: 0.0382172092795372
step: 60, loss: 0.12345796078443527
step: 70, loss: 0.08629650622606277
step: 80, loss: 0.033292241394519806
step: 90, loss: 0.13623034954071045
step: 100, loss: 0.307438462972641
step: 110, loss: 0.14011813700199127
step: 120, loss: 0.09266550093889236
step: 130, loss: 0.11212494224309921
step: 140, loss: 0.34575775265693665
step: 150, loss: 0.09824033826589584
step: 160, loss: 0.21741469204425812

```

```

step: 170, loss: 0.12829914689064026
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7669172932330827, f1=0.7559808612440191,
best_f1=0.7559808612440191
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14016571640968323
step: 10, loss: 0.08018378913402557
step: 20, loss: 0.07740788906812668
step: 30, loss: 0.2578282058238983
step: 40, loss: 0.08566073328256607
step: 50, loss: 0.13336388766765594
step: 60, loss: 0.08021154254674911
step: 70, loss: 0.08590752631425858
step: 80, loss: 0.047036003321409225
step: 90, loss: 0.21178023517131805
step: 100, loss: 0.020496198907494545
step: 110, loss: 0.07938499003648758
step: 120, loss: 0.06333784013986588
step: 130, loss: 0.1425594985485077
step: 140, loss: 0.05002064257860184
step: 150, loss: 0.08955232799053192
step: 160, loss: 0.0494915246963501
step: 170, loss: 0.08185029774904251
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7728459530026109, f1=0.7669172932330827,
best_f1=0.7669172932330827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05426459386944771
step: 10, loss: 0.03372087702155113
step: 20, loss: 0.03139173611998558
step: 30, loss: 0.03151148930191994
step: 40, loss: 0.0025260362308472395

```

```

step: 50, loss: 0.05875542759895325
step: 60, loss: 0.31397444009780884
step: 70, loss: 0.010271042585372925
step: 80, loss: 0.07792916148900986
step: 90, loss: 0.05873329937458038
step: 100, loss: 0.1789478063583374
step: 110, loss: 0.10825372487306595
step: 120, loss: 0.02587568387389183
step: 130, loss: 0.062473833560943604
step: 140, loss: 0.08705305308103561
step: 150, loss: 0.12902668118476868
step: 160, loss: 0.07928363978862762
step: 170, loss: 0.014957894571125507
/usr/local/lib/python3.7/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=0.7609756097560975, f1=0.7703016241299303,
best_f1=0.7669172932330827
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017416128888726234
step: 10, loss: 0.006490418221801519
step: 20, loss: 0.08978545665740967
step: 30, loss: 0.008168577216565609
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.10276439785957336
step: 50, loss: 0.1886325180530548
step: 60, loss: 0.02425079233944416
step: 70, loss: 0.14308419823646545
step: 80, loss: 0.08333475142717361
step: 90, loss: 0.10797087848186493
step: 100, loss: 0.1414894014596939
step: 110, loss: 0.04984355345368385
step: 120, loss: 0.10853646695613861
step: 130, loss: 0.032828979194164276
step: 140, loss: 0.026228327304124832
step: 150, loss: 0.015511252917349339
step: 160, loss: 0.012815392576158047
step: 170, loss: 0.004345873836427927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7938931297709925, f1=0.7837150127226464,

```

```

best_f1=0.7837150127226464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009723008843138814
step: 10, loss: 0.019251661375164986
step: 20, loss: 0.006137113552540541
step: 30, loss: 0.10749093443155289
step: 40, loss: 0.0215646643191576
step: 50, loss: 0.06897779554128647
step: 60, loss: 0.006100276950746775
step: 70, loss: 0.10045921057462692
step: 80, loss: 0.023663099855184555
step: 90, loss: 0.0308522991836071
step: 100, loss: 0.009221049956977367
step: 110, loss: 0.000358502846211195
step: 120, loss: 0.02465413138270378
step: 130, loss: 0.009542613290250301
step: 140, loss: 0.004911658354103565
step: 150, loss: 0.033168941736221313
step: 160, loss: 0.06263312697410583
step: 170, loss: 0.09246088564395905
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7989690721649484, f1=0.7951219512195121,
best_f1=0.7951219512195121
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008702676277607679
step: 10, loss: 0.0033266874961555004
step: 20, loss: 0.012659572064876556
step: 30, loss: 0.017646800726652145
step: 40, loss: 0.008067849092185497
step: 50, loss: 0.00972488708794117
step: 60, loss: 0.00044111296301707625
step: 70, loss: 0.022138865664601326
step: 80, loss: 0.03609864041209221
step: 90, loss: 0.0002526627213228494
step: 100, loss: 0.00041741097811609507
step: 110, loss: 0.02314937114715576
step: 120, loss: 0.0005479935789480805
step: 130, loss: 0.08975660800933838

```

```

step: 140, loss: 0.0051292660646140575
step: 150, loss: 0.0015385339502245188
step: 160, loss: 0.01057574711740017
step: 170, loss: 0.011522735469043255
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8010610079575595, f1=0.7989690721649484,
best_f1=0.7989690721649484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011453295592218637
step: 10, loss: 0.0016012032283470035
step: 20, loss: 0.0007159037631936371
step: 30, loss: 0.019895954057574272
step: 40, loss: 0.00019000381871592253
step: 50, loss: 0.002026304369792342
step: 60, loss: 0.016347579658031464
step: 70, loss: 0.004774716217070818
step: 80, loss: 0.00970317330211401
step: 90, loss: 0.006746246013790369
step: 100, loss: 0.07048908621072769
step: 110, loss: 0.06073266267776489
step: 120, loss: 0.0014939996181055903
step: 130, loss: 0.004906741436570883
step: 140, loss: 0.0012960669118911028
step: 150, loss: 0.0022796031553298235
step: 160, loss: 0.007857694290578365
step: 170, loss: 0.009688693098723888
/usr/local/lib/python3.7/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=0.7513812154696132, f1=0.7417582417582418,
best_f1=0.7989690721649484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005141738336533308
step: 10, loss: 0.0011864654952660203
step: 20, loss: 0.0011080339318141341
step: 30, loss: 0.0016011834377422929
step: 40, loss: 0.0005603623576462269
step: 50, loss: 0.000211476391996257

```

```

step: 60, loss: 0.00485726585611701
step: 70, loss: 0.017608370631933212
step: 80, loss: 0.0011873203329741955
step: 90, loss: 0.020657166838645935
step: 100, loss: 0.04819655790925026
step: 110, loss: 0.002286008559167385
step: 120, loss: 0.0506039559841156
step: 130, loss: 0.1460554450750351
step: 140, loss: 0.005550294648855925
step: 150, loss: 0.027288077399134636
step: 160, loss: 0.011956220492720604
step: 170, loss: 0.0017446192214265466
/usr/local/lib/python3.7/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=0.7878787878787877, f1=0.7709251101321585,
best_f1=0.7989690721649484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04308857023715973
step: 10, loss: 0.00020389420387800783
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 20, loss: 0.15305674076080322
step: 30, loss: 0.004146676044911146
step: 40, loss: 0.0005489125032909214
step: 50, loss: 0.061331965029239655
step: 60, loss: 0.00415191613137722
step: 70, loss: 0.038607459515333176
step: 80, loss: 0.0057866573333740234
step: 90, loss: 0.00724796112626791
step: 100, loss: 0.0009097831207327545
step: 110, loss: 0.0009322210098616779
step: 120, loss: 0.0007600924582220614
step: 130, loss: 0.00047687088954262435
step: 140, loss: 0.004565997049212456
step: 150, loss: 0.0007176477229222655
step: 160, loss: 0.00022217899095267057
step: 170, loss: 0.00024382799165323377
/usr/local/lib/python3.7/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=0.7949367088607596, f1=0.7951219512195121,
best_f1=0.7989690721649484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032749444246292114

```

```

step: 10, loss: 0.01090905349701643
step: 20, loss: 0.0002574566169641912
step: 30, loss: 0.0001221149432240054
step: 40, loss: 0.0011555084493011236
step: 50, loss: 0.0004450917767826468
step: 60, loss: 0.006977875251322985
step: 70, loss: 0.00029323159833438694
step: 80, loss: 0.00018523834296502173
step: 90, loss: 0.0002549227501731366
step: 100, loss: 0.01542644388973713
step: 110, loss: 0.12889285385608673
step: 120, loss: 0.0002906818699557334
step: 130, loss: 0.00015129851817619056
step: 140, loss: 0.00022891572734806687
step: 150, loss: 0.0009157481254078448
step: 160, loss: 0.028967201709747314
step: 170, loss: 0.0008985287859104574
/usr/local/lib/python3.7/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=0.7990074441687345, f1=0.7858823529411765,
best_f1=0.7989690721649484
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005090140271931887
step: 10, loss: 0.00018467227346263826
step: 20, loss: 0.0017438099021092057
step: 30, loss: 0.0006787814782001078
step: 40, loss: 0.001357087749056518
step: 50, loss: 0.0003101710171904415
step: 60, loss: 0.00032214721431955695
step: 70, loss: 0.052136484533548355
step: 80, loss: 0.00026439217617735267
step: 90, loss: 0.000350871414411813
step: 100, loss: 0.0009384783916175365
step: 110, loss: 0.0002468094462528825
step: 120, loss: 0.0036148815415799618
step: 130, loss: 0.1520456075668335
step: 140, loss: 0.0030570372473448515
step: 150, loss: 0.001164968591183424
step: 160, loss: 0.00015078438445925713
step: 170, loss: 0.000629876391030848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8060453400503778, f1=0.7903614457831325,
best_f1=0.7903614457831325
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02551787532866001
step: 10, loss: 0.00047631311463192105
step: 20, loss: 0.000764504715334624
step: 30, loss: 0.00015878125850576907
step: 40, loss: 0.015795540064573288
step: 50, loss: 0.00013028662942815572
step: 60, loss: 0.0022768089547753334
step: 70, loss: 0.0013134053442627192
step: 80, loss: 0.00044818632886745036
step: 90, loss: 9.665441029937938e-05
step: 100, loss: 0.0012917056446895003
step: 110, loss: 8.759079355513677e-05
step: 120, loss: 0.019623521715402603
step: 130, loss: 0.00014472912880592048
step: 140, loss: 0.0007792628020979464
step: 150, loss: 0.05090973153710365
step: 160, loss: 0.0005043313140049577
step: 170, loss: 0.00030007361783646047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8099999999999999, f1=0.7877358490566038,
best_f1=0.7877358490566038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002684962237253785
step: 10, loss: 8.374505705432966e-05
step: 20, loss: 0.00018493911193218082
step: 30, loss: 0.0001718347193673253
step: 40, loss: 0.00013363097968976945
step: 50, loss: 0.00012722726387437433
step: 60, loss: 0.00013587449211627245
step: 70, loss: 0.00022308457118924707
step: 80, loss: 0.00027688153204508126
step: 90, loss: 8.327541581820697e-05
step: 100, loss: 0.0002260479814140126

```



```

step: 110, loss: 0.00021223675867076963
step: 120, loss: 0.028732536360621452
step: 130, loss: 0.007931840606033802
step: 140, loss: 0.0001494777825428173
step: 150, loss: 0.0009601020137779415
step: 160, loss: 0.00015397844254039228
step: 170, loss: 0.00027501489967107773
/usr/local/lib/python3.7/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=0.8051282051282052, f1=0.8029556650246306,
best_f1=0.7877358490566038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033043374423868954
step: 10, loss: 0.0003230276342947036
step: 20, loss: 0.01298128068447113
step: 30, loss: 8.143199374899268e-05
step: 40, loss: 0.00013747200137004256
step: 50, loss: 8.302720380015671e-05
step: 60, loss: 0.022859172895550728
step: 70, loss: 0.0001669505872996524
step: 80, loss: 0.03149697557091713
step: 90, loss: 0.00019729915948119015
step: 100, loss: 0.0002201776806032285
step: 110, loss: 0.0006137315649539232
step: 120, loss: 0.00035085593117401004
step: 130, loss: 0.00011026396532543004
step: 140, loss: 0.00011347964755259454
step: 150, loss: 0.00022467119561042637
step: 160, loss: 0.0003061838506255299
step: 170, loss: 0.0012398915132507682
/usr/local/lib/python3.7/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=0.8062015503875968, f1=0.79, best_f1=0.7877358490566038
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:07, 256.66it/s]
load_f1 = 0.6651053864168619
real_f1 = 0.6728538283062646
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:19, 225.10it/s]

```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```

[11]: # 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 bert-base-uncased \
        --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 bert-base-uncased \
  --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: 0.6126011610031128
step: 10, loss: 0.6527372002601624
step: 20, loss: 0.4880896508693695
step: 30, loss: 0.3975338041782379
step: 40, loss: 0.26567211747169495
step: 50, loss: 0.06396333873271942
step: 60, loss: 0.1458476036787033
step: 70, loss: 0.05747273564338684
step: 80, loss: 0.12282168865203857
step: 90, loss: 0.10568390786647797
step: 100, loss: 0.004571777302771807
step: 110, loss: 0.23206840455532074
step: 120, loss: 0.01871507614850998
step: 130, loss: 0.008458291180431843
step: 140, loss: 0.005634055007249117
step: 150, loss: 0.01165102794766426
step: 160, loss: 0.017222965136170387
step: 170, loss: 0.13111567497253418
step: 180, loss: 0.12118849903345108
step: 190, loss: 0.08691630512475967
step: 200, loss: 0.08075296878814697
step: 210, loss: 0.013369560241699219
step: 220, loss: 0.030246881768107414
step: 230, loss: 0.032949332147836685
```

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

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

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

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=0.9671574178935448, 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: 0.01253865472972393
step: 10, loss: 0.0076027060858905315
step: 20, loss: 0.3110095262527466
step: 30, loss: 0.18061085045337677
step: 40, loss: 0.05921277031302452
step: 50, loss: 0.02766549028456211
```

```

step: 60, loss: 0.007250667084008455
step: 70, loss: 0.030923942103981972
step: 80, loss: 0.20837868750095367
step: 90, loss: 0.04754062741994858
step: 100, loss: 0.14061880111694336
step: 110, loss: 0.20782533288002014
step: 120, loss: 0.11756771057844162
step: 130, loss: 0.07157785445451736
step: 140, loss: 0.0025072877760976553
step: 150, loss: 0.017243552953004837
step: 160, loss: 0.0907628983259201
step: 170, loss: 0.0018413832876831293
step: 180, loss: 0.01760738715529442
step: 190, loss: 0.007904977537691593
step: 200, loss: 0.008219534531235695
step: 210, loss: 0.013042888604104519
step: 220, loss: 0.020889298990368843
step: 230, loss: 0.03653707355260849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9702315325248071, f1=0.9721913236929923,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006974378135055304
step: 10, loss: 0.04768693447113037
step: 20, loss: 0.0033438962418586016
step: 30, loss: 0.002296461956575513
step: 40, loss: 0.12155673652887344
step: 50, loss: 0.054248493164777756
step: 60, loss: 0.022989608347415924
step: 70, loss: 0.0053657954558730125
step: 80, loss: 0.0015065384795889258
step: 90, loss: 0.05937513709068298
step: 100, loss: 0.00046900598681531847
step: 110, loss: 0.0010191226610913873
step: 120, loss: 0.01518743485212326
step: 130, loss: 0.0007851509726606309
step: 140, loss: 0.0011057317024096847
step: 150, loss: 0.04129253327846527
step: 160, loss: 0.031200271099805832
step: 170, loss: 0.02510489523410797

```

```

step: 180, loss: 0.05049625784158707
step: 190, loss: 0.02623898535966873
step: 200, loss: 0.04598816856741905
step: 210, loss: 0.0038255969993770123
step: 220, loss: 0.0017997699324041605
step: 230, loss: 0.0006485473713837564
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9765363128491621, f1=0.967452300785634,
best_f1=0.967452300785634
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040577654726803303
step: 10, loss: 0.01128549687564373
step: 20, loss: 0.0007738117128610611
step: 30, loss: 0.00032089813612401485
step: 40, loss: 0.0016232941998168826
step: 50, loss: 0.0007447195239365101
step: 60, loss: 0.002017645863816142
step: 70, loss: 0.0037231387104839087
step: 80, loss: 0.004729273729026318
step: 90, loss: 0.016417495906352997
step: 100, loss: 0.007113283034414053
step: 110, loss: 0.0033640225883573294
step: 120, loss: 0.037177674472332
step: 130, loss: 0.04056518152356148
step: 140, loss: 0.0021210063714534044
step: 150, loss: 0.17067913711071014
step: 160, loss: 0.06522431969642639
step: 170, loss: 0.024223821237683296
step: 180, loss: 0.00037140343920327723
step: 190, loss: 0.136310875415802
step: 200, loss: 0.002623399253934622
step: 210, loss: 0.06511382013559341
step: 220, loss: 0.0005827300483360887
step: 230, loss: 0.0009018222917802632
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774774774774775, f1=0.9775280898876404,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011228544171899557
step: 10, loss: 0.00043154411832802
step: 20, loss: 0.0067100077867507935
step: 30, loss: 0.0002708283718675375
step: 40, loss: 0.0007661355193704367
step: 50, loss: 0.0029982293490320444
step: 60, loss: 0.2049357295036316
step: 70, loss: 0.0007483907393179834
step: 80, loss: 0.001456500613130629
step: 90, loss: 0.004827443975955248
step: 100, loss: 0.0004902297514490783
step: 110, loss: 0.00041862347279675305
step: 120, loss: 0.00010211296466877684
step: 130, loss: 0.003991322126239538
step: 140, loss: 0.0025546157266944647
step: 150, loss: 0.002892864402383566
step: 160, loss: 0.0005046247388236225
step: 170, loss: 0.002657569944858551
step: 180, loss: 0.004293206613510847
step: 190, loss: 0.04632715880870819
step: 200, loss: 0.0023071381729096174
step: 210, loss: 0.009015209972858429
step: 220, loss: 0.0016860940959304571
step: 230, loss: 0.012870448641479015
/usr/local/lib/python3.7/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=0.9670828603859251, f1=0.9681093394077448,
best_f1=0.9775280898876404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07734512537717819
step: 10, loss: 0.00982637982815504
step: 20, loss: 0.01048331893980503
step: 30, loss: 0.00024192439741455019
step: 40, loss: 0.0003339753020554781
step: 50, loss: 0.03219113126397133
step: 60, loss: 0.00023476807109545916
step: 70, loss: 0.008227691985666752
step: 80, loss: 0.001730539370328188
step: 90, loss: 0.0052446285262703896

```

```

step: 100, loss: 0.04189801216125488
step: 110, loss: 0.0006238609785214067
step: 120, loss: 0.11243776977062225
step: 130, loss: 0.0008317552856169641
step: 140, loss: 0.0013718227855861187
step: 150, loss: 0.010919883847236633
step: 160, loss: 0.0004400262259878218
step: 170, loss: 0.0004257207619957626
step: 180, loss: 0.03459533303976059
step: 190, loss: 0.08390911668539047
step: 200, loss: 0.0007020271150395274
step: 210, loss: 0.0022760231513530016
step: 220, loss: 0.0015270696021616459
step: 230, loss: 0.0331089124083519
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9776286353467561, f1=0.9765363128491621,
best_f1=0.9765363128491621
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026623098528943956
step: 10, loss: 0.000246195966610685
step: 20, loss: 0.00023191932996269315
step: 30, loss: 0.00023768364917486906
step: 40, loss: 0.0003199357306584716
step: 50, loss: 0.0004210088518448174
step: 60, loss: 0.0002999085118062794
step: 70, loss: 0.041103560477495193
step: 80, loss: 0.0006694376352243125
step: 90, loss: 0.00038683009915985167
step: 100, loss: 0.00048388485447503626
step: 110, loss: 0.00023814353335183114
step: 120, loss: 0.00035611240309663117
step: 130, loss: 0.0015442841686308384
step: 140, loss: 0.0011006025597453117
step: 150, loss: 0.00015873955271672457
step: 160, loss: 0.06704866141080856
step: 170, loss: 0.021408287808299065
step: 180, loss: 0.0006800832925364375
step: 190, loss: 0.0014242319157347083
step: 200, loss: 0.02453378587961197
step: 210, loss: 7.828049274394289e-05

```



```

step: 220, loss: 0.00025916993035934865
step: 230, loss: 0.00034151767613366246
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9785310734463276, f1=0.9819004524886877,
best_f1=0.9819004524886877
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002627172798383981
step: 10, loss: 0.00025435510906390846
step: 20, loss: 0.00011701216135406867
step: 30, loss: 0.0006247369456104934
step: 40, loss: 0.0011538820108398795
step: 50, loss: 0.0003870457876473665
step: 60, loss: 0.00020578672410920262
step: 70, loss: 0.0003970218531321734
step: 80, loss: 0.0008485792786814272
step: 90, loss: 7.159342931117862e-05
step: 100, loss: 0.00016901103663258255
step: 110, loss: 0.00011736681335605681
step: 120, loss: 0.00021281266526784748
step: 130, loss: 0.00011421350063756108
step: 140, loss: 7.857051241444424e-05
step: 150, loss: 8.75258119776845e-05
step: 160, loss: 0.00024711398873478174
step: 170, loss: 0.00036551945959217846
step: 180, loss: 0.00010565595584921539
step: 190, loss: 0.000251814752118662
step: 200, loss: 0.00010167754226131365
step: 210, loss: 0.0008236333378590643
step: 220, loss: 0.00012020250869682059
step: 230, loss: 0.00014973471115808934
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808773903262092, f1=0.9785310734463276,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.104819815140218e-05
step: 10, loss: 0.00011664468911476433
step: 20, loss: 0.0003807905304711312
step: 30, loss: 0.00011219624866498634
step: 40, loss: 0.05258239060640335
step: 50, loss: 7.527583511546254e-05
step: 60, loss: 0.0001021713251248002
step: 70, loss: 0.0002750180719885975
step: 80, loss: 7.167068542912602e-05
step: 90, loss: 0.00012131348921684548
step: 100, loss: 7.793032273184508e-05
step: 110, loss: 6.397895776899531e-05
step: 120, loss: 3.518001176416874e-05
step: 130, loss: 3.9746679249219596e-05
step: 140, loss: 6.641366053372622e-05
step: 150, loss: 0.0028258010279387236
step: 160, loss: 8.307098323712125e-05
step: 170, loss: 9.220601350534707e-05
step: 180, loss: 0.0001132795077865012
step: 190, loss: 6.351860065478832e-05
step: 200, loss: 6.877210398670286e-05
step: 210, loss: 5.679506648448296e-05
step: 220, loss: 5.052207416156307e-05
step: 230, loss: 4.886560418526642e-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=0.9776785714285714, f1=0.9743016759776536,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.0140406276332214e-05
step: 10, loss: 8.395867916988209e-05
step: 20, loss: 5.683355266228318e-05
step: 30, loss: 6.294447666732594e-05
step: 40, loss: 4.241314309183508e-05
step: 50, loss: 5.9329158830223605e-05
step: 60, loss: 6.304839916992933e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 70, loss: 8.791083382675424e-05
step: 80, loss: 0.00010611765901558101
step: 90, loss: 6.842521543148905e-05
step: 100, loss: 4.783384065376595e-05
step: 110, loss: 0.0001510446600150317
step: 120, loss: 0.0002228591765742749

```

```

step: 130, loss: 0.0007968784775584936
step: 140, loss: 0.03880402445793152
step: 150, loss: 0.020866425707936287
step: 160, loss: 9.43236009334214e-05
step: 170, loss: 3.2755022402852774e-05
step: 180, loss: 0.001973672304302454
step: 190, loss: 0.000526083167642355
step: 200, loss: 4.6951106924097985e-05
step: 210, loss: 0.00020079135720152408
step: 220, loss: 4.31622720498126e-05
step: 230, loss: 0.0002885922440327704
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9764837625979844,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.845056562568061e-05
step: 10, loss: 4.4902244553668424e-05
step: 20, loss: 0.00011348613043082878
step: 30, loss: 0.00031880565802566707
step: 40, loss: 0.00010918391490122303
step: 50, loss: 3.733322591870092e-05
step: 60, loss: 0.00011218366853427142
step: 70, loss: 3.802524952334352e-05
step: 80, loss: 7.655729132238775e-05
step: 90, loss: 3.234927135054022e-05
step: 100, loss: 6.029678479535505e-05
step: 110, loss: 2.6511946998653002e-05
step: 120, loss: 9.34827548917383e-05
step: 130, loss: 0.0002989622007589787
step: 140, loss: 3.3511052606627345e-05
step: 150, loss: 0.0524294339120388
step: 160, loss: 0.00018534513947088271
step: 170, loss: 0.011144941672682762
step: 180, loss: 0.0005698499735444784
step: 190, loss: 0.00011863037070725113
step: 200, loss: 0.0015379536198452115
step: 210, loss: 2.8743326765834354e-05
step: 220, loss: 4.916566103929654e-05
step: 230, loss: 3.632331208791584e-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=0.9776286353467561, f1=0.9776286353467561,
best_f1=0.9785310734463276

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.036112558329478e-05
step: 10, loss: 4.522918243310414e-05
step: 20, loss: 0.00018791590991895646
step: 30, loss: 0.00015192190767265856
step: 40, loss: 4.5497359678847715e-05
step: 50, loss: 0.010488727129995823
step: 60, loss: 0.005049597471952438
step: 70, loss: 0.020970623940229416
step: 80, loss: 5.3875981393503025e-05
step: 90, loss: 9.371626219945028e-05
step: 100, loss: 2.4407365344814025e-05
step: 110, loss: 3.344792276038788e-05
step: 120, loss: 4.715426257462241e-05
step: 130, loss: 0.00033395891659893095
step: 140, loss: 8.188914944184944e-05
step: 150, loss: 0.00024024331651162356
step: 160, loss: 0.00016157595382537693
step: 170, loss: 7.408853707602248e-05
step: 180, loss: 2.7994552510790527e-05
step: 190, loss: 3.14470635203179e-05
step: 200, loss: 2.0246447093086317e-05
step: 210, loss: 2.9935408747405745e-05
step: 220, loss: 6.117126758908853e-05
step: 230, loss: 0.00015477236593142152
/usr/local/lib/python3.7/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=0.9787709497206705, f1=0.9742441209406495,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016266391321551055
step: 10, loss: 3.845337778329849e-05
step: 20, loss: 2.7283060262561776e-05
step: 30, loss: 0.0002217446599388495
step: 40, loss: 3.541048863553442e-05
step: 50, loss: 0.00014313901192508638
step: 60, loss: 0.0002587540075182915
step: 70, loss: 2.760253300948534e-05
step: 80, loss: 6.31136863376014e-05
step: 90, loss: 0.00014164933236315846
step: 100, loss: 2.169926665374078e-05
step: 110, loss: 0.00966478232294321
step: 120, loss: 5.06639844388701e-05

```

```

step: 130, loss: 3.0274473829194903e-05
step: 140, loss: 0.00020012696040794253
step: 150, loss: 3.9176695281639695e-05
step: 160, loss: 2.4362027033930644e-05
step: 170, loss: 5.7294870202895254e-05
step: 180, loss: 2.98422783089336e-05
step: 190, loss: 2.477236557751894e-05
step: 200, loss: 5.889424937777221e-05
step: 210, loss: 1.6368703654734418e-05
step: 220, loss: 0.0003025256155524403
step: 230, loss: 1.9237035303376615e-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=0.9776286353467561, f1=0.9742441209406495,
best_f1=0.9785310734463276
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000351289170794189
step: 10, loss: 3.670648220577277e-05
step: 20, loss: 2.520096313674003e-05
step: 30, loss: 2.6623985831975006e-05
step: 40, loss: 0.002804488642141223
step: 50, loss: 3.992136043962091e-05
step: 60, loss: 2.7558813599171117e-05
step: 70, loss: 3.102315167780034e-05
step: 80, loss: 4.2487667087698355e-05
step: 90, loss: 5.26755211467389e-05
step: 100, loss: 3.166717942804098e-05
step: 110, loss: 2.701869379961863e-05
step: 120, loss: 2.3736549337627366e-05
step: 130, loss: 2.9246346457512118e-05
step: 140, loss: 3.573883805074729e-05
step: 150, loss: 0.00030269642593339086
step: 160, loss: 0.0022169495932757854
step: 170, loss: 1.3492873222276103e-05
step: 180, loss: 2.3353133656200953e-05
step: 190, loss: 4.8975478421198204e-05
step: 200, loss: 4.236956374370493e-05
step: 210, loss: 0.0009549983660690486
step: 220, loss: 1.8130560420104302e-05
step: 230, loss: 0.0031987898983061314
/usr/local/lib/python3.7/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=0.9787234042553192, f1=0.9763779527559054,
best_f1=0.9785310734463276

```

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

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

```
step: 0, loss: 2.5714356524986215e-05
step: 10, loss: 1.478171816415852e-05
step: 20, loss: 2.857936851796694e-05
step: 30, loss: 3.774502329179086e-05
step: 40, loss: 4.612604243448004e-05
step: 50, loss: 0.006188246421515942
step: 60, loss: 3.6930588976247236e-05
step: 70, loss: 4.206090306979604e-05
step: 80, loss: 2.7554760890780017e-05
step: 90, loss: 3.14032549795229e-05
step: 100, loss: 0.00019554788013920188
step: 110, loss: 2.6560226615401916e-05
step: 120, loss: 3.3041840652003884e-05
step: 130, loss: 3.796105011133477e-05
step: 140, loss: 2.1453399313031696e-05
step: 150, loss: 5.0999860832234845e-05
step: 160, loss: 1.9903747670468874e-05
step: 170, loss: 1.820141915231943e-05
step: 180, loss: 0.00031287307501770556
step: 190, loss: 4.23486708314158e-05
step: 200, loss: 2.758108894340694e-05
step: 210, loss: 6.56067204545252e-05
step: 220, loss: 2.3237635105033405e-05
step: 230, loss: 2.667580520210322e-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=0.9786276715410572, f1=0.9785794813979707,
best_f1=0.9785310734463276
```

```
mv: cannot stat '*_dev.pt': No such file or directory
```

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

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

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

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

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

```
enabled                : True
opt_level              : 02
```

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:11, 211.49it/s]
load_f1 = 0.9765886287625419
real_f1 = 0.9744160177975528
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 [00:17, 252.42it/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 bert-base-uncased \
--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 bert-base-uncased \
  --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: 0.6195212602615356

step: 10, loss: 0.5765916109085083
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.5651190280914307
step: 30, loss: 0.2706012725830078
step: 40, loss: 0.20202946662902832
step: 50, loss: 0.3509480953216553
step: 60, loss: 0.10224489867687225
step: 70, loss: 0.17626447975635529
step: 80, loss: 0.08984517306089401
step: 90, loss: 0.3696531653404236
step: 100, loss: 0.028100015595555305
step: 110, loss: 0.09497007727622986
step: 120, loss: 0.2015731781721115
step: 130, loss: 0.11670743674039841
step: 140, loss: 0.18831464648246765
step: 150, loss: 0.02651393413543701
step: 160, loss: 0.02908453531563282
step: 170, loss: 0.3086468279361725
step: 180, loss: 0.1458704173564911
step: 190, loss: 0.010649560019373894
step: 200, loss: 0.18204112350940704
step: 210, loss: 0.01593373902142048
step: 220, loss: 0.1767669916152954
step: 230, loss: 0.10715630650520325
step: 240, loss: 0.06764689832925797
step: 250, loss: 0.012844357639551163
step: 260, loss: 0.08256451040506363
step: 270, loss: 0.01834874041378498
step: 280, loss: 0.02294364757835865
step: 290, loss: 0.1675165593624115
step: 300, loss: 0.04470326751470566
step: 310, loss: 0.15878064930438995
step: 320, loss: 0.13066436350345612
step: 330, loss: 0.033280279487371445
step: 340, loss: 0.025908589363098145
step: 350, loss: 0.05665979161858559
step: 360, loss: 0.1484445333480835
step: 370, loss: 0.07763749361038208
step: 380, loss: 0.04872081056237221
step: 390, loss: 0.20568925142288208
step: 400, loss: 0.25816407799720764
step: 410, loss: 0.040617093443870544
step: 420, loss: 0.05815301835536957
step: 430, loss: 0.1224735751748085
step: 440, loss: 0.08387583494186401
step: 450, loss: 0.011273612268269062
step: 460, loss: 0.0209552850574255
step: 470, loss: 0.1320737600326538

```

step: 480, loss: 0.032671984285116196
step: 490, loss: 0.05299540236592293
step: 500, loss: 0.08585385233163834
step: 510, loss: 0.16160838305950165
step: 520, loss: 0.07363683730363846
step: 530, loss: 0.004940960556268692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9272811486799445, f1=0.925925925925926,
best_f1=0.925925925925926
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13253583014011383
step: 10, loss: 0.05691390484571457
step: 20, loss: 0.0443231500685215
step: 30, loss: 0.04267364367842674
step: 40, loss: 0.09868130832910538
step: 50, loss: 0.14161553978919983
step: 60, loss: 0.0065262350253760815
step: 70, loss: 0.03372414410114288
step: 80, loss: 0.07566142082214355
step: 90, loss: 0.02701016142964363
step: 100, loss: 0.023448975756764412
step: 110, loss: 0.022619172930717468
step: 120, loss: 0.06318160891532898
step: 130, loss: 0.21218527853488922
step: 140, loss: 0.053803060203790665
step: 150, loss: 0.12056922912597656
step: 160, loss: 0.018872451037168503
step: 170, loss: 0.03600763529539108
step: 180, loss: 0.05060649290680885
step: 190, loss: 0.06778207421302795
step: 200, loss: 0.033464182168245316
step: 210, loss: 0.06493229418992996
step: 220, loss: 0.0512474849820137
step: 230, loss: 0.004148425534367561
step: 240, loss: 0.006506418343633413
step: 250, loss: 0.08905289322137833
step: 260, loss: 0.00426099356263876
step: 270, loss: 0.3049338459968567
step: 280, loss: 0.03438476473093033
step: 290, loss: 0.05466708913445473

```

```

step: 300, loss: 0.15193502604961395
step: 310, loss: 0.022829484194517136
step: 320, loss: 0.07010871171951294
step: 330, loss: 0.039905428886413574
step: 340, loss: 0.05101151019334793
step: 350, loss: 0.01749131642282009
step: 360, loss: 0.07318411022424698
step: 370, loss: 0.09608688205480576
step: 380, loss: 0.056131161749362946
step: 390, loss: 0.0410529188811779
step: 400, loss: 0.03470990061759949
step: 410, loss: 0.004191124811768532
step: 420, loss: 0.17803867161273956
step: 430, loss: 0.00814494676887989
step: 440, loss: 0.21102160215377808
step: 450, loss: 0.024528518319129944
step: 460, loss: 0.11545268446207047
step: 470, loss: 0.04326992854475975
step: 480, loss: 0.3169393837451935
step: 490, loss: 0.026460090652108192
step: 500, loss: 0.14582459628582
step: 510, loss: 0.015128230676054955
step: 520, loss: 0.04868721216917038
step: 530, loss: 0.008245762437582016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9336448598130841, f1=0.9369453526389537,
best_f1=0.9369453526389537
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06975121051073074
step: 10, loss: 0.01457066461443901
step: 20, loss: 0.031811464577913284
step: 30, loss: 0.01792074181139469
step: 40, loss: 0.012826628051698208
step: 50, loss: 0.05442037060856819
step: 60, loss: 0.012984794564545155
step: 70, loss: 0.0013384164776653051
step: 80, loss: 0.0016516183968633413
step: 90, loss: 0.009069678373634815
step: 100, loss: 0.016295913606882095
step: 110, loss: 0.00478900084272027

```

```
step: 120, loss: 0.0029611103236675262
step: 130, loss: 0.013963501900434494
step: 140, loss: 0.10192937403917313
step: 150, loss: 0.0150893060490489
step: 160, loss: 0.011182018555700779
step: 170, loss: 0.06815513223409653
step: 180, loss: 0.08099626004695892
step: 190, loss: 0.009581775404512882
step: 200, loss: 0.057819169014692307
step: 210, loss: 0.044069401919841766
step: 220, loss: 0.03628029674291611
step: 230, loss: 0.0315762534737587
step: 240, loss: 0.0031131517607718706
step: 250, loss: 0.08951357752084732
step: 260, loss: 0.016452694311738014
step: 270, loss: 0.004566090647131205
step: 280, loss: 0.09791824221611023
step: 290, loss: 0.0024751676246523857
step: 300, loss: 0.024491287767887115
step: 310, loss: 0.0008743611397221684
step: 320, loss: 0.09691835194826126
step: 330, loss: 0.013068515807390213
step: 340, loss: 0.018561463803052902
step: 350, loss: 0.006175433751195669
step: 360, loss: 0.05202168598771095
step: 370, loss: 0.012686355970799923
step: 380, loss: 0.01161677110940218
step: 390, loss: 0.12569865584373474
step: 400, loss: 0.02742885798215866
step: 410, loss: 0.005186567083001137
step: 420, loss: 0.37090256810188293
step: 430, loss: 0.011059198528528214
step: 440, loss: 0.002702200785279274
step: 450, loss: 0.06044469401240349
step: 460, loss: 0.05031370371580124
step: 470, loss: 0.012616841122508049
step: 480, loss: 0.002214850392192602
step: 490, loss: 0.007964401505887508
step: 500, loss: 0.049142878502607346
step: 510, loss: 0.009635289199650288
step: 520, loss: 0.19199681282043457
step: 530, loss: 0.16258513927459717
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9365671641791044, f1=0.9285042333019755,
best_f1=0.9285042333019755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007492916192859411
step: 10, loss: 0.002759241499006748
step: 20, loss: 0.0018770701717585325
step: 30, loss: 0.003592473454773426
step: 40, loss: 0.001042340649291873
step: 50, loss: 0.012981999665498734
step: 60, loss: 0.0038510605227202177
step: 70, loss: 0.0014340993948280811
step: 80, loss: 0.016506753861904144
step: 90, loss: 0.07209735363721848
step: 100, loss: 0.006707184482365847
step: 110, loss: 0.004747482016682625
step: 120, loss: 0.0006303103873506188
step: 130, loss: 0.0005276952288113534
step: 140, loss: 0.0009954130509868264
step: 150, loss: 0.03559499979019165
step: 160, loss: 0.002098396187648177
step: 170, loss: 0.0032029643189162016
step: 180, loss: 0.02569214440882206
step: 190, loss: 0.005160018336027861
step: 200, loss: 0.007016859948635101
step: 210, loss: 0.030721500515937805
step: 220, loss: 0.001031847088597715
step: 230, loss: 0.027903635054826736
step: 240, loss: 0.00390934431925416
step: 250, loss: 0.011181836947798729
step: 260, loss: 0.0013633670751005411
step: 270, loss: 0.0036296045873314142
step: 280, loss: 0.021004987880587578
step: 290, loss: 0.018851129338145256
step: 300, loss: 0.0013192385667935014
step: 310, loss: 0.0266936793923378
step: 320, loss: 0.020666565746068954
step: 330, loss: 0.01259284932166338
step: 340, loss: 0.00614632572978735
step: 350, loss: 0.006644006818532944
step: 360, loss: 0.015004978515207767
step: 370, loss: 0.007939326576888561
step: 380, loss: 0.0029079008381813765
step: 390, loss: 0.001929916557855904
step: 400, loss: 0.0020435508340597153
step: 410, loss: 0.0042672534473240376

```

```

step: 420, loss: 0.011159161105751991
step: 430, loss: 0.11848296225070953
step: 440, loss: 0.022369230166077614
step: 450, loss: 0.08038304001092911
step: 460, loss: 0.03354453295469284
step: 470, loss: 0.02126815915107727
step: 480, loss: 0.0034557355102151632
step: 490, loss: 0.0020450856536626816
step: 500, loss: 0.01171035598963499
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 510, loss: 0.10241449624300003
step: 520, loss: 0.22167733311653137
step: 530, loss: 0.01547135878354311
/usr/local/lib/python3.7/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=0.9320565435476517, f1=0.9324137931034482,
best_f1=0.9285042333019755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10346756875514984
step: 10, loss: 0.050809722393751144
step: 20, loss: 0.002394131850451231
step: 30, loss: 0.0027144707273691893
step: 40, loss: 0.003729406278580427
step: 50, loss: 0.0006626932881772518
step: 60, loss: 0.007492128759622574
step: 70, loss: 0.014752635732293129
step: 80, loss: 0.0006109355017542839
step: 90, loss: 0.022680558264255524
step: 100, loss: 0.007785417605191469
step: 110, loss: 0.0007461195345968008
step: 120, loss: 0.012243775650858879
step: 130, loss: 0.00045587419299408793
step: 140, loss: 0.0007731199730187654
step: 150, loss: 0.00953905563801527
step: 160, loss: 0.00036853132769465446
step: 170, loss: 0.05217902362346649
step: 180, loss: 0.0008127790642902255
step: 190, loss: 0.004659904167056084
step: 200, loss: 0.0005628687213174999
step: 210, loss: 0.006715515162795782
step: 220, loss: 0.0037746152374893427
step: 230, loss: 0.01240452378988266
step: 240, loss: 0.004047929774969816
step: 250, loss: 0.013637342490255833
step: 260, loss: 0.012312465347349644

```

```

step: 270, loss: 0.0015669780550524592
step: 280, loss: 0.10568627715110779
step: 290, loss: 0.142738938331604
step: 300, loss: 0.014439794234931469
step: 310, loss: 0.0065361387096345425
step: 320, loss: 0.052671585232019424
step: 330, loss: 0.0983051136136055
step: 340, loss: 0.0018195408629253507
step: 350, loss: 0.0005099042900837958
step: 360, loss: 0.015026210807263851
step: 370, loss: 0.025258265435695648
step: 380, loss: 0.0004290786455385387
step: 390, loss: 0.00019371301459614187
step: 400, loss: 0.0017979086842387915
step: 410, loss: 0.015757808461785316
step: 420, loss: 0.00868980772793293
step: 430, loss: 0.01564270630478859
step: 440, loss: 0.00795766618102789
step: 450, loss: 0.16484849154949188
step: 460, loss: 0.002219551708549261
step: 470, loss: 0.07151418924331665
step: 480, loss: 0.009022342972457409
step: 490, loss: 0.0022501021157950163
step: 500, loss: 0.005111621227115393
step: 510, loss: 0.002897809259593487
step: 520, loss: 0.0005703885108232498
step: 530, loss: 0.0011419864604249597
/usr/local/lib/python3.7/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=0.9297094657919399, f1=0.9279026217228464,
best_f1=0.9285042333019755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00919105764478445
step: 10, loss: 0.00018541251483839005
step: 20, loss: 0.09814789146184921
step: 30, loss: 0.0009171895799227059
step: 40, loss: 0.0641070008277893
step: 50, loss: 0.01043115183711052
step: 60, loss: 0.00034362688893452287
step: 70, loss: 0.0003771652700379491
step: 80, loss: 0.0009246297413483262
step: 90, loss: 0.00017960413242690265
step: 100, loss: 0.0025795435067266226
step: 110, loss: 0.020391609519720078
step: 120, loss: 0.0008681142353452742

```

```

step: 130, loss: 0.0019777962006628513
step: 140, loss: 0.04704009369015694
step: 150, loss: 0.0024491252843290567
step: 160, loss: 0.004819035530090332
step: 170, loss: 0.007534568663686514
step: 180, loss: 0.0034055905416607857
step: 190, loss: 0.00036079654819332063
step: 200, loss: 0.00019711852655746043
step: 210, loss: 0.0029485225677490234
step: 220, loss: 0.00020589676569215953
step: 230, loss: 0.0003828612680081278
step: 240, loss: 0.001260321238078177
step: 250, loss: 0.015686841681599617
step: 260, loss: 9.636517643230036e-05
step: 270, loss: 0.05077638477087021
step: 280, loss: 0.0008388060377910733
step: 290, loss: 0.00031374130048789084
step: 300, loss: 0.0028347070328891277
step: 310, loss: 0.0006088482914492488
step: 320, loss: 0.0009457318228669465
step: 330, loss: 0.00045203379704616964
step: 340, loss: 0.0054770647548139095
step: 350, loss: 0.0006256599444895983
step: 360, loss: 0.0031525942031294107
step: 370, loss: 0.006640706676989794
step: 380, loss: 0.0001342079194728285
step: 390, loss: 0.036135852336883545
step: 400, loss: 0.0001995862985495478
step: 410, loss: 0.0011306841624900699
step: 420, loss: 0.030599620193243027
step: 430, loss: 0.036485761404037476
step: 440, loss: 9.907634375849739e-05
step: 450, loss: 0.0040053571574389935
step: 460, loss: 7.730462675681338e-05
step: 470, loss: 0.0005005723796784878
step: 480, loss: 0.0074946996755898
step: 490, loss: 0.004402755293995142
step: 500, loss: 0.0011524097062647343
step: 510, loss: 0.019221654161810875
step: 520, loss: 0.021230660378932953
step: 530, loss: 0.00446250569075346
/usr/local/lib/python3.7/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=0.9296173830892772, f1=0.9180483183325437,
best_f1=0.9285042333019755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017569514457136393
step: 10, loss: 0.0010619444074109197
step: 20, loss: 0.0008445741259492934
step: 30, loss: 0.0005572860245592892
step: 40, loss: 0.09300735592842102
step: 50, loss: 0.01120565365999937
step: 60, loss: 0.000998001778498292
step: 70, loss: 0.011368019506335258
step: 80, loss: 0.033232927322387695
step: 90, loss: 0.0004457421600818634
step: 100, loss: 0.0003673556202556938
step: 110, loss: 0.002418326446786523
step: 120, loss: 0.0009219153434969485
step: 130, loss: 0.01245032250881195
step: 140, loss: 0.00021481038129422814
step: 150, loss: 0.00023374526062980294
step: 160, loss: 9.034380491357297e-05
step: 170, loss: 0.004382355138659477
step: 180, loss: 0.0005790955037809908
step: 190, loss: 0.002783124800771475
step: 200, loss: 0.0005719860782846808
step: 210, loss: 0.00376767385751009
step: 220, loss: 0.000876014877576381
step: 230, loss: 0.0027522847522050142
step: 240, loss: 0.004722734913229942
step: 250, loss: 0.0006051286472938955
step: 260, loss: 0.0006143860518932343
step: 270, loss: 0.00179457594640553
step: 280, loss: 0.00024736931663937867
step: 290, loss: 0.0004081601509824395
step: 300, loss: 0.0018931833328679204
step: 310, loss: 7.993172766873613e-05
step: 320, loss: 5.9337289712857455e-05
step: 330, loss: 0.00045364152174443007
step: 340, loss: 0.0418863371014595
step: 350, loss: 0.04335104674100876
step: 360, loss: 0.0016367494827136397
step: 370, loss: 0.0001066683980752714
step: 380, loss: 0.1734425574541092
step: 390, loss: 0.016422223299741745
step: 400, loss: 0.001609547296538949
step: 410, loss: 0.001306032296270132
step: 420, loss: 0.0013005140936002135
step: 430, loss: 0.0001222880237037316
step: 440, loss: 0.02634650282561779
step: 450, loss: 0.000767081102821976
step: 460, loss: 0.0004013872821815312
```

```

step: 470, loss: 0.003984385170042515
step: 480, loss: 0.06565458327531815
step: 490, loss: 0.0007386918878182769
step: 500, loss: 0.027139397338032722
step: 510, loss: 0.003992523066699505
step: 520, loss: 0.0014857195783406496
step: 530, loss: 0.0014126661699265242
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9402985074626865, f1=0.9274004683840749,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005649511702358723
step: 10, loss: 0.02952253259718418
step: 20, loss: 0.0008161460282281041
step: 30, loss: 0.002377950819209218
step: 40, loss: 0.0002004967536777258
step: 50, loss: 0.0008602454327046871
step: 60, loss: 0.0025440293829888105
step: 70, loss: 0.00021612024283967912
step: 80, loss: 0.0001906217512441799
step: 90, loss: 0.00016565168334636837
step: 100, loss: 0.0005826420383527875
step: 110, loss: 0.0007577879587188363
step: 120, loss: 0.0003309389285277575
step: 130, loss: 0.003623978467658162
step: 140, loss: 0.0013498942134901881
step: 150, loss: 0.020085854455828667
step: 160, loss: 0.00012922476162202656
step: 170, loss: 0.001724899746477604
step: 180, loss: 4.489791172090918e-05
step: 190, loss: 0.0002774458553176373
step: 200, loss: 0.002653789008036256
step: 210, loss: 0.0005975327803753316
step: 220, loss: 0.0007945027900859714
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 230, loss: 0.01891220360994339
step: 240, loss: 7.476117025362328e-05
step: 250, loss: 0.04799959063529968
step: 260, loss: 0.00019383961625862867
step: 270, loss: 0.010745150037109852

```

```

step: 280, loss: 0.0013496900210157037
step: 290, loss: 0.00019909509865101427
step: 300, loss: 0.0007668319740332663
step: 310, loss: 0.00016569055151194334
step: 320, loss: 0.019122913479804993
step: 330, loss: 0.001248071901500225
step: 340, loss: 0.0017471954924985766
step: 350, loss: 0.0002674379211384803
step: 360, loss: 9.704499825602397e-05
step: 370, loss: 0.0006114374264143407
step: 380, loss: 0.0003256397321820259
step: 390, loss: 0.09835092723369598
step: 400, loss: 0.0008269482059404254
step: 410, loss: 0.0005140863941051066
step: 420, loss: 0.0003349227481521666
step: 430, loss: 0.014690757729113102
step: 440, loss: 0.0004658898978959769
step: 450, loss: 0.00036529835779219866
step: 460, loss: 0.004712519235908985
step: 470, loss: 4.791836909134872e-05
step: 480, loss: 0.004296299535781145
step: 490, loss: 0.00012910933583043516
step: 500, loss: 9.6085415862035e-05
step: 510, loss: 0.00012685115507338196
step: 520, loss: 0.002945579355582595
step: 530, loss: 2.8113478037994355e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 8: dev_f1=0.9332087809434843, f1=0.924953095684803,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001299262628890574
step: 10, loss: 7.051140710245818e-05
step: 20, loss: 0.043456826359033585
step: 30, loss: 4.146090941503644e-05
step: 40, loss: 0.00048336584586650133
step: 50, loss: 3.14868739224039e-05
step: 60, loss: 8.85341432876885e-05
step: 70, loss: 0.0013906779931858182
step: 80, loss: 0.005386752542108297
step: 90, loss: 0.00010001431655837223
step: 100, loss: 6.605195812880993e-05
step: 110, loss: 5.2356150263221934e-05
step: 120, loss: 5.911454718443565e-05
step: 130, loss: 0.002570055890828371

```

```

step: 140, loss: 0.0016758347628638148
step: 150, loss: 3.9567326894029975e-05
step: 160, loss: 0.00020641941227950156
step: 170, loss: 0.004567799158394337
step: 180, loss: 7.228335016407073e-05
step: 190, loss: 3.566546365618706e-05
step: 200, loss: 5.2656665502581745e-05
step: 210, loss: 2.2857735530124046e-05
step: 220, loss: 6.460276199504733e-05
step: 230, loss: 3.087332152063027e-05
step: 240, loss: 6.97513751219958e-05
step: 250, loss: 0.00035152232157997787
step: 260, loss: 0.009947218000888824
step: 270, loss: 0.002002053428441286
step: 280, loss: 0.0006545149371959269
step: 290, loss: 0.023777266964316368
step: 300, loss: 3.691590245580301e-05
step: 310, loss: 0.0008561965660192072
step: 320, loss: 0.029801707714796066
step: 330, loss: 8.99302976904437e-05
step: 340, loss: 0.00015198663459159434
step: 350, loss: 0.0015037282137200236
step: 360, loss: 5.536551907425746e-05
step: 370, loss: 0.039273813366889954
step: 380, loss: 0.0036603461485356092
step: 390, loss: 3.8934660551603884e-05
step: 400, loss: 0.00015583647473249584
step: 410, loss: 0.0007578007644042373
step: 420, loss: 0.00013494912127498537
step: 430, loss: 5.213002805248834e-05
step: 440, loss: 0.001938357949256897
step: 450, loss: 0.00010485580423846841
step: 460, loss: 0.00016464856162201613
step: 470, loss: 3.153271973133087e-05
step: 480, loss: 4.082572922925465e-05
step: 490, loss: 0.0005395914195105433
step: 500, loss: 0.005132363177835941
step: 510, loss: 0.0005518170073628426
step: 520, loss: 0.002375206211581826
step: 530, loss: 8.784775855019689e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 4096.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9371215649743828, f1=0.9302107728337237,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007114681648090482
step: 10, loss: 7.716967957094312e-05
step: 20, loss: 5.131991201778874e-05
step: 30, loss: 0.00043628033017739654
step: 40, loss: 0.0003010965301655233
step: 50, loss: 7.060759526211768e-05
step: 60, loss: 3.8357342418748885e-05
step: 70, loss: 2.440757270960603e-05
step: 80, loss: 3.082543844357133e-05
step: 90, loss: 3.119808388873935e-05
step: 100, loss: 0.0030369299929589033
step: 110, loss: 1.926676486618817e-05
step: 120, loss: 0.0001581173128215596
step: 130, loss: 0.00011900313256774098
step: 140, loss: 0.0023177252151072025
step: 150, loss: 3.1286901503335685e-05
step: 160, loss: 2.681357364053838e-05
step: 170, loss: 0.00011587983317440376
step: 180, loss: 5.8231718867318705e-05
step: 190, loss: 2.4910445063142106e-05
step: 200, loss: 0.0002955813833978027
step: 210, loss: 0.0004172890039626509
step: 220, loss: 0.00043394090607762337
step: 230, loss: 0.0009984101634472609
step: 240, loss: 0.0008375747711397707
step: 250, loss: 2.1501895389519632e-05
step: 260, loss: 0.0019073731964454055
step: 270, loss: 0.00010120368096977472
step: 280, loss: 6.008896525599994e-05
step: 290, loss: 2.1758813090855256e-05
step: 300, loss: 2.1151585315237753e-05
step: 310, loss: 0.0001575176283949986
step: 320, loss: 0.0006395849050022662
step: 330, loss: 3.642858428065665e-05
step: 340, loss: 3.4458924346836284e-05
step: 350, loss: 3.781215855269693e-05
step: 360, loss: 6.036788909113966e-05
step: 370, loss: 0.00036748923594132066
step: 380, loss: 0.00011972471838817
step: 390, loss: 0.002573252422735095
step: 400, loss: 0.00020737577870022506
step: 410, loss: 0.0007487852708436549
step: 420, loss: 0.0003101800393778831
step: 430, loss: 1.7497368389740586e-05
step: 440, loss: 8.550158963771537e-05
step: 450, loss: 3.198936246917583e-05
step: 460, loss: 2.2335994799504988e-05
```

```

step: 470, loss: 0.0057886624708771706
step: 480, loss: 3.692874815897085e-05
step: 490, loss: 0.048236824572086334
step: 500, loss: 0.033361632376909256
step: 510, loss: 0.006759658921509981
step: 520, loss: 0.0003125386137980968
step: 530, loss: 0.0001384680945193395
/usr/local/lib/python3.7/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=0.9345266132830901, f1=0.9243065350258581,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00041772244730964303
step: 10, loss: 0.0015536301070824265
step: 20, loss: 0.000367827364243567
step: 30, loss: 0.0010490705026313663
step: 40, loss: 2.4675597160239704e-05
step: 50, loss: 0.0003986067895311862
step: 60, loss: 0.03640146926045418
step: 70, loss: 3.3316926419502124e-05
step: 80, loss: 3.165148154948838e-05
step: 90, loss: 7.792739052092656e-05
step: 100, loss: 0.00013950321590527892
step: 110, loss: 0.0029623592272400856
step: 120, loss: 6.747170118615031e-05
step: 130, loss: 1.9616807549027726e-05
step: 140, loss: 0.001663219416514039
step: 150, loss: 1.817902193579357e-05
step: 160, loss: 0.00029436047771014273
step: 170, loss: 2.6441397494636476e-05
step: 180, loss: 3.981712143286131e-05
step: 190, loss: 0.0005939934053458273
step: 200, loss: 0.06585351377725601
step: 210, loss: 0.00012789653555955738
step: 220, loss: 0.005029933527112007
step: 230, loss: 2.3460472220904194e-05
step: 240, loss: 0.001699427724815905
step: 250, loss: 2.6462377718416974e-05
step: 260, loss: 8.456937939627096e-05
step: 270, loss: 0.004472623113542795
step: 280, loss: 8.276801236206666e-05
step: 290, loss: 5.5196756875375286e-05
step: 300, loss: 3.50514892488718e-05
step: 310, loss: 2.4172342818928882e-05
step: 320, loss: 4.2951200157403946e-05

```

```

step: 330, loss: 6.021286390023306e-05
step: 340, loss: 0.0015518140280619264
step: 350, loss: 0.004899853374809027
step: 360, loss: 2.365064574405551e-05
step: 370, loss: 0.09552344679832458
step: 380, loss: 2.0723340639960952e-05
step: 390, loss: 3.827585169347003e-05
step: 400, loss: 4.5884287828812376e-05
step: 410, loss: 0.00036606189678423107
step: 420, loss: 4.5309938286663964e-05
step: 430, loss: 0.00017957072122953832
step: 440, loss: 0.00013035045412834734
step: 450, loss: 3.748282324522734e-05
step: 460, loss: 0.00013507110998034477
step: 470, loss: 0.0014946083538234234
step: 480, loss: 0.00046258026850409806
step: 490, loss: 4.6745171857764944e-05
step: 500, loss: 0.00021093030227348208
step: 510, loss: 0.00011351694411132485
step: 520, loss: 3.0479235647362657e-05
step: 530, loss: 0.0004043527296744287
/usr/local/lib/python3.7/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=0.9311969839773798, f1=0.9227144203581528,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.5705975682940334e-05
step: 10, loss: 1.7635233234614134e-05
step: 20, loss: 0.0019430659012869
step: 30, loss: 3.9394915802404284e-05
step: 40, loss: 8.253785199485719e-05
step: 50, loss: 0.03316566348075867
step: 60, loss: 0.0007463502115570009
step: 70, loss: 2.8545975510496646e-05
step: 80, loss: 3.5097546060569584e-05
step: 90, loss: 2.6637677365215495e-05
step: 100, loss: 2.9346099836402573e-05
step: 110, loss: 3.214443495380692e-05
step: 120, loss: 4.434544825926423e-05
step: 130, loss: 5.3667557949665934e-05
step: 140, loss: 2.038434467976913e-05
step: 150, loss: 1.4677462786494289e-05
step: 160, loss: 1.5906758562778123e-05
step: 170, loss: 0.00014764147636014968
step: 180, loss: 2.614285949675832e-05

```

```

step: 190, loss: 3.473127435427159e-05
step: 200, loss: 2.8027865482727066e-05
step: 210, loss: 3.318592644063756e-05
step: 220, loss: 0.0006201990763656795
step: 230, loss: 1.174202134279767e-05
step: 240, loss: 3.466839189059101e-05
step: 250, loss: 1.9706216335180216e-05
step: 260, loss: 5.120319110574201e-05
step: 270, loss: 2.7197684175916947e-05
step: 280, loss: 0.0016430671093985438
step: 290, loss: 0.0006626711110584438
step: 300, loss: 0.00013850329560227692
step: 310, loss: 0.00010544699034653604
step: 320, loss: 1.9739847630262375e-05
step: 330, loss: 5.188323848415166e-05
step: 340, loss: 0.020327774807810783
step: 350, loss: 2.939539081125986e-05
step: 360, loss: 0.0006560836918652058
step: 370, loss: 0.0010972294257953763
step: 380, loss: 0.00010496489994693547
step: 390, loss: 9.683080861577764e-05
step: 400, loss: 6.834600935690105e-05
step: 410, loss: 0.0007855519070290029
step: 420, loss: 0.006071730982512236
step: 430, loss: 0.00023687808425165713
step: 440, loss: 0.0002494214568287134
step: 450, loss: 5.9113528550369665e-05
step: 460, loss: 0.0003947076038457453
step: 470, loss: 3.3747823181329295e-05
step: 480, loss: 4.328229624661617e-05
step: 490, loss: 0.00014057981024961919
step: 500, loss: 0.00031248919549398124
step: 510, loss: 0.00036529178032651544
step: 520, loss: 0.00014420950901694596
step: 530, loss: 6.621011561946943e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9359925788497218, f1=0.9290681502086231,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006201095529831946
step: 10, loss: 3.187154288752936e-05
step: 20, loss: 0.0006850226200185716
step: 30, loss: 7.673575601074845e-05
step: 40, loss: 2.744991797953844e-05

```


step: 50, loss: 0.0007504629320465028
step: 60, loss: 1.728871211525984e-05
step: 70, loss: 3.12221709464211e-05
step: 80, loss: 3.818076947936788e-05
step: 90, loss: 0.016168231144547462
step: 100, loss: 0.0032151085324585438
step: 110, loss: 1.4629058568971232e-05
step: 120, loss: 2.4105527700157836e-05
step: 130, loss: 1.6290485291392542e-05
step: 140, loss: 2.6943243938148953e-05
step: 150, loss: 0.00010614572238409892
step: 160, loss: 1.782880826795008e-05
step: 170, loss: 9.289410809287801e-05
step: 180, loss: 2.3397922632284462e-05
step: 190, loss: 0.0006416430696845055
step: 200, loss: 4.204725701129064e-05
step: 210, loss: 3.05758039758075e-05
step: 220, loss: 1.978429099835921e-05
step: 230, loss: 3.62702994607389e-05
step: 240, loss: 0.0010473011061549187
step: 250, loss: 1.8998433006345294e-05
step: 260, loss: 2.9415055905701593e-05
step: 270, loss: 2.9060405722702853e-05
step: 280, loss: 1.7869897419586778e-05
step: 290, loss: 2.8170101359137334e-05
step: 300, loss: 2.4478144041495398e-05
step: 310, loss: 0.0015843127621337771
step: 320, loss: 0.0010837140725925565
step: 330, loss: 0.00034626887645572424
step: 340, loss: 3.410403951420449e-05
step: 350, loss: 2.557227708166465e-05
step: 360, loss: 1.4658866348327138e-05
step: 370, loss: 0.00014241307508200407
step: 380, loss: 9.817010868573561e-05
step: 390, loss: 0.0004302476882003248
step: 400, loss: 4.0936625737231225e-05
step: 410, loss: 0.0002654832205735147
step: 420, loss: 2.2037998860469088e-05
step: 430, loss: 2.1631929485010915e-05
step: 440, loss: 3.0764451366849244e-05
step: 450, loss: 1.788464214769192e-05
step: 460, loss: 0.00021421707060653716
step: 470, loss: 2.467705235176254e-05
step: 480, loss: 3.349994221935049e-05
step: 490, loss: 3.124537761323154e-05
step: 500, loss: 2.8276388547965325e-05
step: 510, loss: 0.0002221287868451327
step: 520, loss: 7.448783435393125e-05

```

step: 530, loss: 1.235667059518164e-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=0.9363891487371375, f1=0.9271028037383178,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007328671053983271
step: 10, loss: 2.3744252757751383e-05
step: 20, loss: 1.4576862668036483e-05
step: 30, loss: 5.477751255966723e-05
step: 40, loss: 0.0002036787336692214
step: 50, loss: 0.0005668638623319566
step: 60, loss: 1.4975248632254079e-05
step: 70, loss: 2.2182746761245653e-05
step: 80, loss: 0.0022441574838012457
step: 90, loss: 0.00017792658763937652
step: 100, loss: 0.0003266620624344796
step: 110, loss: 0.0015486186603084207
step: 120, loss: 0.0001350520469713956
step: 130, loss: 0.021363994106650352
step: 140, loss: 0.0006570083787664771
step: 150, loss: 1.9582970708142966e-05
step: 160, loss: 8.283081115223467e-05
step: 170, loss: 1.8447224647388794e-05
step: 180, loss: 1.2889329809695482e-05
step: 190, loss: 1.4897223991283681e-05
step: 200, loss: 8.817889465717599e-05
step: 210, loss: 1.3671680790139362e-05
step: 220, loss: 7.320961594814435e-05
step: 230, loss: 0.0002610245137475431
step: 240, loss: 0.00011487940355436876
step: 250, loss: 3.771199772018008e-05
step: 260, loss: 0.00020289747044444084
step: 270, loss: 9.505728667136282e-05
step: 280, loss: 1.10081527964212e-05
step: 290, loss: 0.00010937628394458443
step: 300, loss: 0.003583427518606186
step: 310, loss: 2.6750083634397015e-05
step: 320, loss: 0.0008289710385724902
step: 330, loss: 7.503705273848027e-05
step: 340, loss: 5.608183710137382e-05
step: 350, loss: 6.457961717387661e-05
step: 360, loss: 9.652004519011825e-05
step: 370, loss: 3.640705836005509e-05
step: 380, loss: 2.610169212857727e-05

```

```

step: 390, loss: 0.022483158856630325
step: 400, loss: 4.254343730281107e-05
step: 410, loss: 2.2577956769964658e-05
step: 420, loss: 1.3362430763663724e-05
step: 430, loss: 2.34536528296303e-05
step: 440, loss: 4.001690831501037e-05
step: 450, loss: 4.2735471652122214e-05
step: 460, loss: 0.000137261493364349
step: 470, loss: 3.3677672035992146e-05
step: 480, loss: 1.774636075424496e-05
step: 490, loss: 1.3626943655253854e-05
step: 500, loss: 1.0773443136713468e-05
step: 510, loss: 2.412403773632832e-05
step: 520, loss: 4.7541230742353946e-05
step: 530, loss: 0.006902511231601238
/usr/local/lib/python3.7/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=0.9353932584269663, f1=0.9243697478991597,
best_f1=0.9274004683840749
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.2688208698818926e-05
step: 10, loss: 0.0004908169503323734
step: 20, loss: 5.6036049500107765e-05
step: 30, loss: 0.0005577130359597504
step: 40, loss: 0.0014463234692811966
step: 50, loss: 0.00013508116535376757
step: 60, loss: 1.1127347534056753e-05
step: 70, loss: 0.00014487304724752903
step: 80, loss: 2.0462033717194572e-05
step: 90, loss: 0.00017333406140096486
step: 100, loss: 3.130375262117013e-05
step: 110, loss: 2.390286863374058e-05
step: 120, loss: 3.5558263334678486e-05
step: 130, loss: 0.08448746800422668
step: 140, loss: 0.0002297312894370407
step: 150, loss: 2.3009850337984972e-05
step: 160, loss: 5.360241266316734e-05
step: 170, loss: 1.906897705339361e-05
step: 180, loss: 1.2688205060840119e-05
step: 190, loss: 3.868569547194056e-05
step: 200, loss: 3.1839299481362104e-05
step: 210, loss: 0.00013505203241948038
step: 220, loss: 1.4420361367228907e-05
step: 230, loss: 8.547288598492742e-05
step: 240, loss: 1.8528960936237127e-05

```

```

step: 250, loss: 8.475377399008721e-05
step: 260, loss: 1.445395355403889e-05
step: 270, loss: 1.3567360838351306e-05
step: 280, loss: 0.00011186140909558162
step: 290, loss: 9.25728181755403e-06
step: 300, loss: 1.3269265764392912e-05
step: 310, loss: 3.081019895034842e-05
step: 320, loss: 2.5159504730254412e-05
step: 330, loss: 5.5133110436145216e-05
step: 340, loss: 6.559021130669862e-05
step: 350, loss: 2.0661782400566153e-05
step: 360, loss: 0.0013796997955068946
step: 370, loss: 1.8360920876148157e-05
step: 380, loss: 1.443525525246514e-05
step: 390, loss: 2.68652420345461e-05
step: 400, loss: 0.0002444024430587888
step: 410, loss: 0.00021050851501058787
step: 420, loss: 2.253223465231713e-05
step: 430, loss: 4.050443021696992e-05
step: 440, loss: 4.0849812648957595e-05
step: 450, loss: 1.3507632502296474e-05
step: 460, loss: 3.8963127735769376e-05
step: 470, loss: 6.549627869389951e-05
step: 480, loss: 3.252438546041958e-05
step: 490, loss: 1.2259786672075279e-05
step: 500, loss: 1.4204324543243274e-05
step: 510, loss: 3.122250200249255e-05
step: 520, loss: 1.1622805686783977e-05
step: 530, loss: 5.172895544092171e-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=0.9393656716417911, f1=0.9291338582677167,
best_f1=0.9274004683840749
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 [00:30, 187.81it/s]
load_f1 = 0.9367789570835257
real_f1 = 0.9343880874825501
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:23, 185.89it/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 bert-base-uncased \
--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 bert-base-uncased \
  --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: 0.5511002540588379
```

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

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

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

```
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=0.5, f1=0.4210526315789474, best_f1=0.4210526315789474
```

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

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

```
step: 0, loss: 0.49595898389816284
```

```
/usr/local/lib/python3.7/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=0.5, f1=0.43137254901960786, best_f1=0.4210526315789474
```

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

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

```
step: 0, loss: 0.608761727809906
```

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

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

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

```
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=0.6153846153846153, f1=0.45, best_f1=0.45
```

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

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

```
step: 0, loss: 0.25722289085388184
```

```
/usr/local/lib/python3.7/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=0.588235294117647, f1=0.46511627906976755, best_f1=0.45
```

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

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

```
step: 0, loss: 0.2601982355117798
```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7586206896551724, 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: 0.28989189863204956
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.6250000000000001,
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: 0.07463992387056351
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8235294117647058, f1=0.6, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01574951969087124
/usr/local/lib/python3.7/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=0.8125000000000001, f1=0.6486486486486486, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004100667778402567
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.6923076923076924, f1=0.6451612903225806, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```



```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03464384377002716
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.5714285714285714, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006523944437503815
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.5555555555555556, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022012202069163322
/usr/local/lib/python3.7/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=0.75, f1=0.588235294117647, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007865325547754765
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 13: dev_f1=0.75, f1=0.588235294117647, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01728222891688347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.75, f1=0.588235294117647, best_f1=0.6
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00888928771018982
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.75, f1=0.588235294117647, best_f1=0.6
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, 140686.20it/s]

/usr/local/lib/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 = 0.7333333333333334

real_f1 = 0.6875000000000001

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 [00:17, 251.52it/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 bert-base-uncased \
--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 bert-base-uncased \
  --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: 0.6061214804649353
step: 10, loss: 0.605292022228241
step: 20, loss: 0.3887988030910492
step: 30, loss: 0.16632089018821716
step: 40, loss: 0.2980862259864807
step: 50, loss: 0.023008793592453003
step: 60, loss: 0.07408381253480911
step: 70, loss: 0.043123237788677216
step: 80, loss: 0.05447756126523018
step: 90, loss: 0.08592504262924194
step: 100, loss: 0.06027046963572502
step: 110, loss: 0.1549503654241562
step: 120, loss: 0.04348963126540184
step: 130, loss: 0.010631104931235313
step: 140, loss: 0.006415170151740313
step: 150, loss: 0.005066617392003536
step: 160, loss: 0.031125571578741074
step: 170, loss: 0.2472187727689743
step: 180, loss: 0.02043755166232586
step: 190, loss: 0.05109798163175583
step: 200, loss: 0.007822808809578419
step: 210, loss: 0.011585514061152935
step: 220, loss: 0.003178322222083807
step: 230, loss: 0.06561879068613052
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.9799107142857142, f1=0.9772209567198178,

```

```

best_f1=0.9772209567198178
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00391040463000536
step: 10, loss: 0.0073185451328754425
step: 20, loss: 0.1593630611896515
step: 30, loss: 0.19522452354431152
step: 40, loss: 0.07574471086263657
step: 50, loss: 0.002732407534494996
step: 60, loss: 0.005066982004791498
step: 70, loss: 0.02658754214644432
step: 80, loss: 0.0027237138710916042
step: 90, loss: 0.042431578040122986
step: 100, loss: 0.04207243397831917
step: 110, loss: 0.08846672624349594
step: 120, loss: 0.002397255739197135
step: 130, loss: 0.007656731642782688
step: 140, loss: 0.007012381684035063
step: 150, loss: 0.002294846810400486
step: 160, loss: 0.0028200275264680386
step: 170, loss: 0.003613975364714861
step: 180, loss: 0.0037960782647132874
step: 190, loss: 0.10597337037324905
step: 200, loss: 0.02133975736796856
step: 210, loss: 0.0013057009782642126
step: 220, loss: 0.10195448249578476
step: 230, loss: 0.05652402341365814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9898989898989898, f1=0.9886877828054299,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025497782044112682
step: 10, loss: 0.007293289992958307
step: 20, loss: 0.0007460599881596863
step: 30, loss: 0.005540187004953623
step: 40, loss: 0.061241935938596725
step: 50, loss: 0.010928564704954624
step: 60, loss: 0.001921224407851696
step: 70, loss: 0.0015334610361605883

```

```

step: 80, loss: 0.0022647767327725887
step: 90, loss: 0.0438607819378376
step: 100, loss: 0.01093264203518629
step: 110, loss: 0.0022617941722273827
step: 120, loss: 0.015088518150150776
step: 130, loss: 0.002209986327216029
step: 140, loss: 0.0558207631111145
step: 150, loss: 0.004774031694978476
step: 160, loss: 0.003969457000494003
step: 170, loss: 0.015985697507858276
step: 180, loss: 0.0677664577960968
step: 190, loss: 0.034740906208753586
step: 200, loss: 0.008106322027742863
step: 210, loss: 0.05339689180254936
step: 220, loss: 0.0010800090385600924
step: 230, loss: 0.004561154637485743
/usr/local/lib/python3.7/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=0.9898305084745763, f1=0.9853438556933484,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0012332913465797901
step: 10, loss: 0.0004992689937353134
step: 20, loss: 0.00168679712805897
step: 30, loss: 0.001181506086140871
step: 40, loss: 0.003037896938621998
step: 50, loss: 0.0004470893763937056
step: 60, loss: 0.0015111371176317334
step: 70, loss: 0.023020479828119278
step: 80, loss: 0.0032547004520893097
step: 90, loss: 0.029678937047719955
step: 100, loss: 0.0003864556201733649
step: 110, loss: 0.0005500918487086892
step: 120, loss: 0.002838586922734976
step: 130, loss: 0.0007012824644334614
step: 140, loss: 0.0009238818311132491
step: 150, loss: 0.08191327005624771
step: 160, loss: 0.0523848831653595
step: 170, loss: 0.00799059122800827
step: 180, loss: 0.00031192589085549116
step: 190, loss: 0.0013671924825757742
step: 200, loss: 0.000588251743465662
step: 210, loss: 0.007289730943739414
step: 220, loss: 0.0006517054280266166
step: 230, loss: 0.02521209977567196

```

```

/usr/local/lib/python3.7/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=0.9853107344632768, f1=0.9841628959276018,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003307521343231201
step: 10, loss: 0.0008297530584968626
step: 20, loss: 0.04524804651737213
step: 30, loss: 0.0006034502875991166
step: 40, loss: 0.00032416798057965934
step: 50, loss: 0.00018161571642849594
step: 60, loss: 0.022050585597753525
step: 70, loss: 0.00024709911667741835
step: 80, loss: 0.0002273024438181892
step: 90, loss: 0.0007357497233897448
step: 100, loss: 0.0006166977109387517
step: 110, loss: 0.00018078717403113842
step: 120, loss: 0.00016613602929282933
step: 130, loss: 0.005447994451969862
step: 140, loss: 0.01188534963876009
step: 150, loss: 0.0012088310904800892
step: 160, loss: 0.0015741604147478938
step: 170, loss: 0.006835451815277338
step: 180, loss: 0.0021578832529485226
step: 190, loss: 0.001614658278413117
step: 200, loss: 0.024035369977355003
step: 210, loss: 0.04846566542983055
step: 220, loss: 0.002510604914277792
step: 230, loss: 0.00017924320127349347
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9797752808988766,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026651102234609425
step: 10, loss: 0.0008472128538414836
step: 20, loss: 0.016094785183668137
step: 30, loss: 0.001243659295141697
step: 40, loss: 0.0003955017600674182
step: 50, loss: 0.1546647995710373
step: 60, loss: 0.001267924322746694
step: 70, loss: 0.007335457485169172

```

```

step: 80, loss: 0.0066327317617833614
step: 90, loss: 0.008627158589661121
step: 100, loss: 0.003794071264564991
step: 110, loss: 0.0013208953896537423
step: 120, loss: 0.0070229824632406235
step: 130, loss: 0.0024461434222757816
step: 140, loss: 0.0008551275823265314
step: 150, loss: 0.0020201208535581827
step: 160, loss: 0.020362215116620064
step: 170, loss: 0.06317215412855148
step: 180, loss: 0.0048872968181967735
step: 190, loss: 0.03498687967658043
step: 200, loss: 0.0008415665361098945
step: 210, loss: 0.001979719614610076
step: 220, loss: 0.00032641933648847044
step: 230, loss: 0.08591833710670471
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.983277591973244, f1=0.9777777777777777,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063486103899776936
step: 10, loss: 0.00015621005150023848
step: 20, loss: 0.0016572794411331415
step: 30, loss: 0.00028028644737787545
step: 40, loss: 0.0006096633733250201
step: 50, loss: 0.00042008605669252574
step: 60, loss: 0.002174234250560403
step: 70, loss: 0.006069842725992203
step: 80, loss: 0.0007989198202267289
step: 90, loss: 6.856678373878822e-05
step: 100, loss: 0.00013652330380864441
step: 110, loss: 0.001472583506256342
step: 120, loss: 0.0028884713537991047
step: 130, loss: 0.00016076255997177213
step: 140, loss: 0.00016602959658484906
step: 150, loss: 0.0022138634230941534
step: 160, loss: 0.08988130837678909
step: 170, loss: 0.012266106903553009
step: 180, loss: 0.0046605998650193214
step: 190, loss: 0.0002744227822404355
step: 200, loss: 0.15015734732151031
step: 210, loss: 0.00016668302123434842
step: 220, loss: 0.03487051650881767
step: 230, loss: 0.000783730938564986

```



```

/usr/local/lib/python3.7/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=0.987598647125141, f1=0.9784824462061155,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004009282565675676
step: 10, loss: 0.013619153760373592
step: 20, loss: 0.0006676649209111929
step: 30, loss: 0.00018351402832195163
step: 40, loss: 0.007190602831542492
step: 50, loss: 0.000538757536560297
step: 60, loss: 0.0021897819824516773
step: 70, loss: 0.0003764888097066432
step: 80, loss: 0.0022143975365906954
step: 90, loss: 0.00010306006879545748
step: 100, loss: 0.00019591959426179528
step: 110, loss: 0.06084541603922844
step: 120, loss: 0.03827289119362831
step: 130, loss: 0.0006776733789592981
step: 140, loss: 0.0002913026255555451
step: 150, loss: 0.00017509111785329878
step: 160, loss: 0.0003512180701363832
step: 170, loss: 0.00012320633686613292
step: 180, loss: 0.0048442864790558815
step: 190, loss: 0.006660020910203457
step: 200, loss: 0.00025401570019312203
step: 210, loss: 0.0005727014504373074
step: 220, loss: 0.0009546215878799558
step: 230, loss: 0.00020079759997315705
/usr/local/lib/python3.7/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=0.9886363636363636, f1=0.9748283752860413,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002505742886569351
step: 10, loss: 0.001198692829348147
step: 20, loss: 0.03054552525281906
step: 30, loss: 0.0003315930371172726
step: 40, loss: 0.02607487142086029
step: 50, loss: 0.00014509368338622153
step: 60, loss: 0.0002333331503905356
step: 70, loss: 0.00775337778031826

```

```

step: 80, loss: 0.00203350349329412
step: 90, loss: 0.00022285079467110336
step: 100, loss: 0.000362630991730839
step: 110, loss: 0.0004350449307821691
step: 120, loss: 0.00017157854745164514
step: 130, loss: 0.00011129422637168318
step: 140, loss: 0.00010287732584401965
step: 150, loss: 0.001124995294958353
step: 160, loss: 5.669135498465039e-05
step: 170, loss: 0.004693076945841312
step: 180, loss: 0.002948983572423458
step: 190, loss: 9.952845721272752e-05
step: 200, loss: 7.881074998294935e-05
step: 210, loss: 0.006957856006920338
step: 220, loss: 5.794260141556151e-05
step: 230, loss: 5.9679648984456435e-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=0.9886621315192743, f1=0.976054732041049,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.630556501680985e-05
step: 10, loss: 5.3797150030732155e-05
step: 20, loss: 8.624919428257272e-05
step: 30, loss: 8.956748206401244e-05
step: 40, loss: 0.0025615228805691004
step: 50, loss: 5.0073143938789144e-05
step: 60, loss: 0.0002981189463753253
step: 70, loss: 0.0024106719065457582
step: 80, loss: 7.692477811360732e-05
step: 90, loss: 0.0002377850323682651
step: 100, loss: 0.0001148953742813319
step: 110, loss: 0.002005124231800437
step: 120, loss: 0.0001029735867632553
step: 130, loss: 0.0002428300358587876
step: 140, loss: 0.08006144315004349
step: 150, loss: 0.009457878768444061
step: 160, loss: 3.8385001971619204e-05
step: 170, loss: 0.0001840379263740033
step: 180, loss: 0.0003892892855219543
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 190, loss: 0.027955716475844383
step: 200, loss: 0.00010780720913317055
step: 210, loss: 7.331249071285129e-05
step: 220, loss: 7.889342668931931e-05

```

```

step: 230, loss: 0.00014576272224076092
/usr/local/lib/python3.7/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=0.9854423292273236, f1=0.9842342342342343,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.94942825450562e-05
step: 10, loss: 4.829626777791418e-05
step: 20, loss: 0.005543518345803022
step: 30, loss: 0.02466147020459175
step: 40, loss: 7.91358106653206e-05
step: 50, loss: 4.4656066165771335e-05
step: 60, loss: 0.00012588189565576613
step: 70, loss: 0.011390983127057552
step: 80, loss: 4.043734588776715e-05
step: 90, loss: 3.4985350794158876e-05
step: 100, loss: 4.804369746125303e-05
step: 110, loss: 0.00638565793633461
step: 120, loss: 0.0003467602946329862
step: 130, loss: 2.7941558073507622e-05
step: 140, loss: 3.873564492096193e-05
step: 150, loss: 0.05022372677922249
step: 160, loss: 4.1329833038616925e-05
step: 170, loss: 0.04627276584506035
step: 180, loss: 0.010073814541101456
step: 190, loss: 4.204840661259368e-05
step: 200, loss: 0.0011993184452876449
step: 210, loss: 9.398067049914971e-05
step: 220, loss: 0.00020860797667410225
step: 230, loss: 0.0002323703229194507
/usr/local/lib/python3.7/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=0.9854096520763187, f1=0.9831649831649831,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.550706843379885e-05
step: 10, loss: 4.0469774830853567e-05
step: 20, loss: 4.946575791109353e-05
step: 30, loss: 0.00019820136367343366
step: 40, loss: 0.0012995335273444653
step: 50, loss: 0.0003881447482854128
step: 60, loss: 0.042375560849905014

```

```

step: 70, loss: 0.00010929997370112687
step: 80, loss: 5.6391632824670523e-05
step: 90, loss: 5.465991489472799e-05
step: 100, loss: 3.217429912183434e-05
step: 110, loss: 0.00010366107017034665
step: 120, loss: 0.0012216410832479596
step: 130, loss: 4.9161586503032595e-05
step: 140, loss: 0.025181086733937263
step: 150, loss: 8.936645463109016e-05
step: 160, loss: 7.85778829595074e-05
step: 170, loss: 4.834727224078961e-05
step: 180, loss: 0.03303677216172218
step: 190, loss: 0.00021505822951439768
step: 200, loss: 0.000146169972140342
step: 210, loss: 0.0015411594649776816
step: 220, loss: 5.5773623898858204e-05
step: 230, loss: 0.06061113625764847
/usr/local/lib/python3.7/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=0.9876265466816648, f1=0.9785794813979707,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.831220475258306e-05
step: 10, loss: 0.00020445795962587
step: 20, loss: 0.0006657004123553634
step: 30, loss: 0.00020624930039048195
step: 40, loss: 5.382890594773926e-05
step: 50, loss: 8.791712753009051e-05
step: 60, loss: 0.0003833768132608384
step: 70, loss: 2.7789694286184385e-05
step: 80, loss: 0.0004144649428781122
step: 90, loss: 0.000751481216866523
step: 100, loss: 8.39322674437426e-05
step: 110, loss: 0.05114414915442467
step: 120, loss: 0.034279439598321915
step: 130, loss: 0.0001112412428483367
step: 140, loss: 9.8348522442393e-05
step: 150, loss: 5.4620191804133356e-05
step: 160, loss: 0.0015103663317859173
step: 170, loss: 0.00025574929895810783
step: 180, loss: 0.00010180105891777202
step: 190, loss: 4.761349919135682e-05
step: 200, loss: 8.781092765275389e-05
step: 210, loss: 0.00013125441910233349
step: 220, loss: 4.518569039646536e-05

```

```

step: 230, loss: 3.705236304085702e-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=0.9887387387387387, f1=0.9808342728297633,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.8713871870422736e-05
step: 10, loss: 0.009286812506616116
step: 20, loss: 0.001073975465260446
step: 30, loss: 7.467172690667212e-05
step: 40, loss: 0.00013494603626895696
step: 50, loss: 4.706923573394306e-05
step: 60, loss: 3.7976555177010596e-05
step: 70, loss: 3.8518010114785284e-05
step: 80, loss: 2.8721309718093835e-05
step: 90, loss: 3.999659020337276e-05
step: 100, loss: 4.948130663251504e-05
step: 110, loss: 0.0014766567619517446
step: 120, loss: 2.137150841008406e-05
step: 130, loss: 4.347791036707349e-05
step: 140, loss: 0.006006760522723198
step: 150, loss: 6.976364238653332e-05
step: 160, loss: 0.008893095888197422
step: 170, loss: 4.1311068343929946e-05
step: 180, loss: 4.365385393612087e-05
step: 190, loss: 2.9484674087143503e-05
step: 200, loss: 2.736872738751117e-05
step: 210, loss: 5.688942110282369e-05
step: 220, loss: 3.676188862300478e-05
step: 230, loss: 7.563867984572425e-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=0.9887892376681614, f1=0.9808773903262092,
best_f1=0.9886877828054299
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.280575608892832e-05
step: 10, loss: 4.610396717907861e-05
step: 20, loss: 4.39639225078281e-05
step: 30, loss: 0.00011784490197896957
step: 40, loss: 9.268586291000247e-05
step: 50, loss: 0.04390988498926163
step: 60, loss: 3.7195131881162524e-05

```

```

step: 70, loss: 3.9269751141546294e-05
step: 80, loss: 7.330129301408306e-05
step: 90, loss: 0.00023964133288245648
step: 100, loss: 7.353709952440113e-05
step: 110, loss: 0.005154613871127367
step: 120, loss: 5.202697866479866e-05
step: 130, loss: 0.023003092035651207
step: 140, loss: 2.2686623196932487e-05
step: 150, loss: 0.0002849488810170442
step: 160, loss: 2.4448389012832195e-05
step: 170, loss: 2.461601616232656e-05
step: 180, loss: 0.0009564972715452313
step: 190, loss: 0.0008533722721040249
step: 200, loss: 7.669292244827375e-05
step: 210, loss: 0.00024848085013218224
step: 220, loss: 3.7325600715121254e-05
step: 230, loss: 0.003706247778609395
/usr/local/lib/python3.7/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=0.9887892376681614, f1=0.9819819819819819,
best_f1=0.9886877828054299
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, 196.60it/s]
load_f1 = 0.9898989898989898
real_f1 = 0.9865771812080537
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 [00:18, 233.96it/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 bert-base-uncased \
--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 bert-base-uncased \
--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: 0.6285241842269897
```

```
step: 10, loss: 0.5669483542442322
```

```
step: 20, loss: 0.5321688055992126
```

```
step: 30, loss: 0.182898610830307
```

```
step: 40, loss: 0.12531262636184692
```

```
step: 50, loss: 0.2879013121128082
```

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

```
step: 60, loss: 0.13791784644126892
```

```
step: 70, loss: 0.24558648467063904
```



```
step: 80, loss: 0.10841207206249237
step: 90, loss: 0.442481130361557
step: 100, loss: 0.041496392339468
step: 110, loss: 0.05641983449459076
step: 120, loss: 0.12259756028652191
step: 130, loss: 0.10075687617063522
step: 140, loss: 0.08051954209804535
step: 150, loss: 0.05108888819813728
step: 160, loss: 0.022994542494416237
step: 170, loss: 0.286937415599823
step: 180, loss: 0.03054133802652359
step: 190, loss: 0.022461673244833946
step: 200, loss: 0.19407887756824493
step: 210, loss: 0.08699078857898712
step: 220, loss: 0.3465065658092499
step: 230, loss: 0.17822730541229248
step: 240, loss: 0.04034729674458504
step: 250, loss: 0.014584777876734734
step: 260, loss: 0.17137333750724792
step: 270, loss: 0.017274249345064163
step: 280, loss: 0.04840749874711037
step: 290, loss: 0.19885630905628204
step: 300, loss: 0.06790231913328171
step: 310, loss: 0.23013971745967865
step: 320, loss: 0.032489608973264694
step: 330, loss: 0.05417736992239952
step: 340, loss: 0.04506872594356537
step: 350, loss: 0.0610090047121048
step: 360, loss: 0.03434844687581062
step: 370, loss: 0.12791739404201508
step: 380, loss: 0.030965063720941544
step: 390, loss: 0.21496383845806122
step: 400, loss: 0.28014153242111206
step: 410, loss: 0.04271762818098068
step: 420, loss: 0.0114183584228158
step: 430, loss: 0.21828754246234894
step: 440, loss: 0.01379118300974369
step: 450, loss: 0.029481206089258194
step: 460, loss: 0.003077238332480192
step: 470, loss: 0.1630936563014984
step: 480, loss: 0.07224290072917938
step: 490, loss: 0.04520146921277046
step: 500, loss: 0.26821866631507874
step: 510, loss: 0.022634845227003098
step: 520, loss: 0.054193586111068726
step: 530, loss: 0.007056616246700287
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9266943291839558, f1=0.9242843951985227,
best_f1=0.9242843951985227
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21890942752361298
step: 10, loss: 0.09178178012371063
step: 20, loss: 0.01624663732945919
step: 30, loss: 0.023961266502738
step: 40, loss: 0.054806794971227646
step: 50, loss: 0.14817525446414948
step: 60, loss: 0.07633794844150543
step: 70, loss: 0.040345754474401474
step: 80, loss: 0.06655189394950867
step: 90, loss: 0.0688859298825264
step: 100, loss: 0.050431232899427414
step: 110, loss: 0.04060802981257439
step: 120, loss: 0.033867355436086655
step: 130, loss: 0.02649732679128647
step: 140, loss: 0.01918405294418335
step: 150, loss: 0.10421093553304672
step: 160, loss: 0.019495142623782158
step: 170, loss: 0.02913626655936241
step: 180, loss: 0.06749176234006882
step: 190, loss: 0.18134573101997375
step: 200, loss: 0.017483815550804138
step: 210, loss: 0.04889163002371788
step: 220, loss: 0.02062462829053402
step: 230, loss: 0.0050751748494803905
step: 240, loss: 0.11864247173070908
step: 250, loss: 0.0904751792550087
step: 260, loss: 0.0034614368341863155
step: 270, loss: 0.21632522344589233
step: 280, loss: 0.07202326506376266
step: 290, loss: 0.019947391003370285
step: 300, loss: 0.14072732627391815
step: 310, loss: 0.010599079541862011
step: 320, loss: 0.2201213538646698
step: 330, loss: 0.016609760001301765
step: 340, loss: 0.03205070644617081
step: 350, loss: 0.005467991344630718
step: 360, loss: 0.10394211858510971
step: 370, loss: 0.18022724986076355

```

```

step: 380, loss: 0.03132566064596176
step: 390, loss: 0.09874425828456879
step: 400, loss: 0.04208733141422272
step: 410, loss: 0.05810854956507683
step: 420, loss: 0.03329041972756386
step: 430, loss: 0.1611957997083664
step: 440, loss: 0.08732837438583374
step: 450, loss: 0.06489990651607513
step: 460, loss: 0.014789998531341553
step: 470, loss: 0.02895546518266201
step: 480, loss: 0.221199631690979
step: 490, loss: 0.017964059486985207
step: 500, loss: 0.38394248485565186
step: 510, loss: 0.0753404051065445
step: 520, loss: 0.08333936333656311
step: 530, loss: 0.031297262758016586
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9340866290018832, f1=0.9259606373008434,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0901007205247879
step: 10, loss: 0.007743304595351219
step: 20, loss: 0.1424281895160675
step: 30, loss: 0.20885543525218964
step: 40, loss: 0.010421739891171455
step: 50, loss: 0.08587846904993057
step: 60, loss: 0.04031241685152054
step: 70, loss: 0.013980228453874588
step: 80, loss: 0.0014190416550263762
step: 90, loss: 0.010147071443498135
step: 100, loss: 0.02767310105264187
step: 110, loss: 0.007886149920523167
step: 120, loss: 0.025072989985346794
step: 130, loss: 0.010768481530249119
step: 140, loss: 0.052200645208358765
step: 150, loss: 0.03257005289196968
step: 160, loss: 0.005490054842084646
step: 170, loss: 0.18782250583171844
step: 180, loss: 0.010421882383525372
step: 190, loss: 0.06333430856466293

```

```

step: 200, loss: 0.01690562628209591
step: 210, loss: 0.08667097240686417
step: 220, loss: 0.053848862648010254
step: 230, loss: 0.10317162424325943
step: 240, loss: 0.021438395604491234
step: 250, loss: 0.00952671468257904
step: 260, loss: 0.01854466088116169
step: 270, loss: 0.0009400804410688579
step: 280, loss: 0.10425731539726257
step: 290, loss: 0.002325496869161725
step: 300, loss: 0.034709278494119644
step: 310, loss: 0.029337970539927483
step: 320, loss: 0.04080408439040184
step: 330, loss: 0.009390410035848618
step: 340, loss: 0.012159531936049461
step: 350, loss: 0.05510813742876053
step: 360, loss: 0.05192332714796066
step: 370, loss: 0.008980154059827328
step: 380, loss: 0.03944670408964157
step: 390, loss: 0.07472562044858932
step: 400, loss: 0.030650990083813667
step: 410, loss: 0.00524608138948679
step: 420, loss: 0.3432892858982086
step: 430, loss: 0.021345002576708794
step: 440, loss: 0.0065876180306077
step: 450, loss: 0.06876009702682495
step: 460, loss: 0.011559983715415001
step: 470, loss: 0.00883571244776249
step: 480, loss: 0.006713760085403919
step: 490, loss: 0.002779174828901887
step: 500, loss: 0.0025358840357512236
step: 510, loss: 0.011169695295393467
step: 520, loss: 0.10765612125396729
step: 530, loss: 0.18099495768547058
/usr/local/lib/python3.7/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=0.9330210772833724, f1=0.9224952741020794,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010597443208098412
step: 10, loss: 0.01061151921749115
step: 20, loss: 0.0012612675782293081
step: 30, loss: 0.009953909553587437
step: 40, loss: 0.022030729800462723
step: 50, loss: 0.014665858820080757

```

step: 60, loss: 0.010553190484642982
step: 70, loss: 0.046124521642923355
step: 80, loss: 0.02653161622583866
step: 90, loss: 0.025320952758193016
step: 100, loss: 0.0025408302899450064
step: 110, loss: 0.06781080365180969
step: 120, loss: 0.009577454067766666
step: 130, loss: 0.0006381465937010944
step: 140, loss: 0.0033846115693449974
step: 150, loss: 0.017153959721326828
step: 160, loss: 0.1195702850818634
step: 170, loss: 0.013470826670527458
step: 180, loss: 0.00861178245395422
step: 190, loss: 0.036302562803030014
step: 200, loss: 0.002886081114411354
step: 210, loss: 0.07508957386016846
step: 220, loss: 0.009870847687125206
step: 230, loss: 0.10002961754798889
step: 240, loss: 0.006746762432157993
step: 250, loss: 0.014447485096752644
step: 260, loss: 0.001814359100535512
step: 270, loss: 0.01645423285663128
step: 280, loss: 0.03876518830657005
step: 290, loss: 0.11685206741094589
step: 300, loss: 0.0006228065467439592
step: 310, loss: 0.008843693882226944
step: 320, loss: 0.05231742188334465
step: 330, loss: 0.07296323031187057
step: 340, loss: 0.003555557457730174
step: 350, loss: 0.0962429940700531
step: 360, loss: 0.10720962285995483
step: 370, loss: 0.0014056236250326037
step: 380, loss: 0.006823668256402016
step: 390, loss: 0.015712542459368706
step: 400, loss: 0.0035394199658185244
step: 410, loss: 0.018351275473833084
step: 420, loss: 0.01563842222094536
step: 430, loss: 0.1520972102880478
step: 440, loss: 0.011323941871523857
step: 450, loss: 0.020411644130945206
step: 460, loss: 0.00214054505340755
step: 470, loss: 0.01915382593870163
step: 480, loss: 0.1363624930381775
step: 490, loss: 0.0018768602749332786
step: 500, loss: 0.003373140702024102
step: 510, loss: 0.009937630966305733
step: 520, loss: 0.06129147857427597
step: 530, loss: 0.03141492232680321

```
/usr/local/lib/python3.7/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=0.924020764511562, f1=0.9168241965973536,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1024293377995491
step: 10, loss: 0.07605648785829544
step: 20, loss: 0.004116054158657789
step: 30, loss: 0.006748526357114315
step: 40, loss: 0.08719682693481445
step: 50, loss: 0.004854109138250351
step: 60, loss: 0.025629926472902298
step: 70, loss: 0.012865501455962658
step: 80, loss: 0.00047477419138886034
step: 90, loss: 0.016734343022108078
step: 100, loss: 0.009951993823051453
step: 110, loss: 0.0004214121145196259
step: 120, loss: 0.011270679533481598
step: 130, loss: 0.0008638472645543516
step: 140, loss: 0.006962520536035299
step: 150, loss: 0.006797345820814371
step: 160, loss: 0.0006607159157283604
step: 170, loss: 0.028976093977689743
step: 180, loss: 0.003888531122356653
step: 190, loss: 0.006451841443777084
step: 200, loss: 0.00041757934377528727
step: 210, loss: 0.00793563760817051
step: 220, loss: 0.0012075494742020965
step: 230, loss: 0.026978803798556328
step: 240, loss: 0.0001585463178344071
step: 250, loss: 0.0012221974320709705
step: 260, loss: 0.0010746954940259457
step: 270, loss: 0.003408782882615924
step: 280, loss: 0.008576802909374237
step: 290, loss: 0.14384891092777252
step: 300, loss: 0.001784113934263587
step: 310, loss: 0.027619412168860435
step: 320, loss: 0.038975320756435394
step: 330, loss: 0.004174084402620792
step: 340, loss: 0.0026087448932230473
step: 350, loss: 0.002501856070011854
step: 360, loss: 0.0012863914016634226
step: 370, loss: 0.07395047694444656
step: 380, loss: 0.0010647941380739212
step: 390, loss: 0.0005337253678590059
```

```

step: 400, loss: 0.003887779312208295
step: 410, loss: 0.012095445767045021
step: 420, loss: 0.003568827174603939
step: 430, loss: 0.007913751527667046
step: 440, loss: 0.03764210641384125
step: 450, loss: 0.023904047906398773
step: 460, loss: 0.005666851531714201
step: 470, loss: 0.019130764529109
step: 480, loss: 0.0029463900718837976
step: 490, loss: 0.0009380700648762286
step: 500, loss: 0.06466781347990036
step: 510, loss: 0.032108109444379807
step: 520, loss: 0.03586322441697121
step: 530, loss: 0.006639121100306511
/usr/local/lib/python3.7/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=0.9216809933142313, f1=0.9136276391554704,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0174103956669569
step: 10, loss: 0.0009057187708094716
step: 20, loss: 0.08570417761802673
step: 30, loss: 0.001736402977257967
step: 40, loss: 0.04374256730079651
step: 50, loss: 0.03429804742336273
step: 60, loss: 0.007929792627692223
step: 70, loss: 0.002558852545917034
step: 80, loss: 0.11404943466186523
step: 90, loss: 0.00184503267519176
step: 100, loss: 0.0007906618411652744
step: 110, loss: 0.002884302055463195
step: 120, loss: 0.02446235530078411
step: 130, loss: 0.008159087039530277
step: 140, loss: 0.042132988572120667
step: 150, loss: 0.0008723121718503535
step: 160, loss: 0.003754575038328767
step: 170, loss: 0.0028301344718784094
step: 180, loss: 0.0005283191567286849
step: 190, loss: 0.0006278474465943873
step: 200, loss: 0.0002059510734397918
step: 210, loss: 0.04158264398574829
step: 220, loss: 0.002106573898345232
step: 230, loss: 0.0024091065861284733
step: 240, loss: 0.02666548453271389
step: 250, loss: 0.001126629300415516

```

```

step: 260, loss: 0.023524673655629158
step: 270, loss: 0.006907726638019085
step: 280, loss: 0.019943049177527428
step: 290, loss: 0.0012704723048955202
step: 300, loss: 0.007142620626837015
step: 310, loss: 0.0004975608899258077
step: 320, loss: 0.00010721310536609963
step: 330, loss: 0.0004258513217791915
step: 340, loss: 0.0491255559027195
step: 350, loss: 0.008708065375685692
step: 360, loss: 0.023698318749666214
step: 370, loss: 0.08254417032003403
step: 380, loss: 0.0002311818825546652
step: 390, loss: 0.016368167474865913
step: 400, loss: 0.12295406311750412
step: 410, loss: 0.002678483724594116
step: 420, loss: 0.0014474482741206884
step: 430, loss: 0.002560056746006012
step: 440, loss: 0.00012522548786364496
step: 450, loss: 0.00012949023221153766
step: 460, loss: 9.927034261636436e-05
step: 470, loss: 0.00024880500859580934
step: 480, loss: 0.001607969286851585
step: 490, loss: 0.006829915102571249
step: 500, loss: 0.001741070649586618
step: 510, loss: 0.003489800961688161
step: 520, loss: 0.0031919165048748255
step: 530, loss: 0.01896718703210354
/usr/local/lib/python3.7/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=0.926899187768753, f1=0.9111969111969113,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010913291946053505
step: 10, loss: 0.13358764350414276
step: 20, loss: 0.0014406624250113964
step: 30, loss: 0.00033813450136221945
step: 40, loss: 0.005796332843601704
step: 50, loss: 0.01748610846698284
step: 60, loss: 0.01694188080728054
step: 70, loss: 0.128782257437706
step: 80, loss: 0.003024990437552333
step: 90, loss: 0.0015754456399008632
step: 100, loss: 0.0006517560686916113
step: 110, loss: 0.008721916936337948

```



```
step: 120, loss: 0.00985005684196949
step: 130, loss: 0.005548038985580206
step: 140, loss: 0.000570224947296083
step: 150, loss: 0.002194309374317527
step: 160, loss: 0.0004893750883638859
step: 170, loss: 0.0029488529544323683
step: 180, loss: 0.002283995971083641
step: 190, loss: 0.025638682767748833
step: 200, loss: 0.0011331313289701939
step: 210, loss: 0.0020304929930716753
step: 220, loss: 0.0008479313110001385
step: 230, loss: 0.010329591110348701
step: 240, loss: 0.0022011827677488327
step: 250, loss: 0.0010404642671346664
step: 260, loss: 0.000472038023872301
step: 270, loss: 0.09571455419063568
step: 280, loss: 0.0016380869783461094
step: 290, loss: 0.0021948348730802536
step: 300, loss: 0.0015808402094990015
step: 310, loss: 0.004581376910209656
step: 320, loss: 0.00015474454266950488
step: 330, loss: 0.00032037432538345456
step: 340, loss: 0.00018974218983203173
step: 350, loss: 0.0001083182287402451
step: 360, loss: 0.0016623842529952526
step: 370, loss: 0.03054589405655861
step: 380, loss: 0.000904663116671145
step: 390, loss: 0.0003186353715136647
step: 400, loss: 0.015137066133320332
step: 410, loss: 0.0014563470613211393
step: 420, loss: 0.007358611561357975
step: 430, loss: 0.00023000950750429183
step: 440, loss: 0.00020768145623151213
step: 450, loss: 0.0002830904850270599
step: 460, loss: 0.0012508329236879945
step: 470, loss: 0.0025227274745702744
step: 480, loss: 0.11222493648529053
step: 490, loss: 0.00023093802155926824
step: 500, loss: 0.0013438710011541843
step: 510, loss: 0.006117869168519974
step: 520, loss: 0.011779936961829662
step: 530, loss: 0.018000874668359756
/usr/local/lib/python3.7/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=0.9269427640763147, f1=0.9157747789669615,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0009947809157893062
step: 10, loss: 0.0031282047275453806
step: 20, loss: 0.00022325565805658698
step: 30, loss: 0.14457732439041138
step: 40, loss: 0.00021733118046540767
step: 50, loss: 0.029642567038536072
step: 60, loss: 0.0008593470556661487
step: 70, loss: 6.65746774757281e-05
step: 80, loss: 0.0037284495774656534
step: 90, loss: 0.001791840884834528
step: 100, loss: 0.00021240924252197146
step: 110, loss: 0.00037680543027818203
step: 120, loss: 0.0062010991387069225
step: 130, loss: 0.2757044732570648
step: 140, loss: 0.0034825284965336323
step: 150, loss: 0.0008171197841875255
step: 160, loss: 0.07543494552373886
step: 170, loss: 0.013763826340436935
step: 180, loss: 0.0010542075615376234
step: 190, loss: 0.0028873677365481853
step: 200, loss: 0.005004914477467537
step: 210, loss: 0.004525922704488039
step: 220, loss: 0.001545432605780661
step: 230, loss: 0.0049820393323898315
step: 240, loss: 0.013950377702713013
step: 250, loss: 0.00021483012824319303
step: 260, loss: 0.00017371847934555262
step: 270, loss: 0.03850511088967323
step: 280, loss: 0.030442040413618088
step: 290, loss: 0.0074782343581318855
step: 300, loss: 0.0013930853456258774
step: 310, loss: 0.0004371624963823706
step: 320, loss: 0.06029972806572914
step: 330, loss: 0.05665672942996025
step: 340, loss: 0.00011700025061145425
step: 350, loss: 0.00018125917995348573
step: 360, loss: 0.006680522114038467
step: 370, loss: 0.0002860456588678062
step: 380, loss: 0.0013105300022289157
step: 390, loss: 0.12026407569646835
step: 400, loss: 0.00016246414452325553
step: 410, loss: 0.0013779107248410583
step: 420, loss: 0.0076371910981833935
step: 430, loss: 0.0022507179528474808
step: 440, loss: 0.027125095948576927
step: 450, loss: 7.159994856920093e-05
```

Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 460, loss: 0.0004250780330039561
step: 470, loss: 0.00010989073780365288
step: 480, loss: 4.932649608235806e-05
step: 490, loss: 0.0008219322771765292
step: 500, loss: 0.0004139734955970198
step: 510, loss: 0.00029413672746159136
step: 520, loss: 0.000811391684692353
step: 530, loss: 0.046387478709220886
/usr/local/lib/python3.7/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=0.9257403189066059, f1=0.9195820081781009,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01727769896388054
step: 10, loss: 0.0055893100798130035
step: 20, loss: 0.06181266903877258
step: 30, loss: 0.00036199475289322436
step: 40, loss: 0.0007858893368393183
step: 50, loss: 0.0009573093266226351
step: 60, loss: 7.624673889949918e-05
step: 70, loss: 0.04977573826909065
step: 80, loss: 0.00047571983304806054
step: 90, loss: 0.00021133161499164999
step: 100, loss: 0.0004715489922091365
step: 110, loss: 8.133990195346996e-05
step: 120, loss: 0.012172365561127663
step: 130, loss: 0.005729194264858961
step: 140, loss: 8.37109619169496e-05
step: 150, loss: 0.0002237223816337064
step: 160, loss: 0.001177011290565133
step: 170, loss: 0.0028899405151605606
step: 180, loss: 0.00012057839921908453
step: 190, loss: 4.5756645704386756e-05
step: 200, loss: 0.00032746707438491285
step: 210, loss: 0.0152514623478055
step: 220, loss: 0.00038821931229904294
step: 230, loss: 4.454405279830098e-05
step: 240, loss: 0.033083848655223846
step: 250, loss: 0.001851867069490254
step: 260, loss: 0.00929450336843729
step: 270, loss: 6.675560143776238e-05
step: 280, loss: 0.0012676535407081246
step: 290, loss: 0.010704873129725456
step: 300, loss: 0.00014513457426801324

```

step: 310, loss: 0.0062699648551642895
step: 320, loss: 0.05494087561964989
step: 330, loss: 0.0013933792943134904
step: 340, loss: 0.0010654510697349906
step: 350, loss: 0.06977984309196472
step: 360, loss: 0.00034631136804819107
step: 370, loss: 0.005054487381130457
step: 380, loss: 0.0007754673133604228
step: 390, loss: 0.00033541175071150064
step: 400, loss: 0.0010332047240808606
step: 410, loss: 0.0006520762108266354
step: 420, loss: 0.0015778725501149893
step: 430, loss: 0.0007339830044656992
step: 440, loss: 0.0018763345433399081
step: 450, loss: 0.00021478504640981555
step: 460, loss: 0.0001697464322205633
step: 470, loss: 4.770298983203247e-05
step: 480, loss: 0.001222248189151287
step: 490, loss: 0.0006111955735832453
step: 500, loss: 0.0032193525694310665
step: 510, loss: 0.0004404279461596161
step: 520, loss: 0.000606061308644712
step: 530, loss: 0.00011073880887124687
/usr/local/lib/python3.7/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=0.9279486002753558, f1=0.9227230910763569,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012299054651521146
step: 10, loss: 0.00027717521879822016
step: 20, loss: 6.432789086829871e-05
step: 30, loss: 0.00044266143231652677
step: 40, loss: 6.700262019876391e-05
step: 50, loss: 0.0007456768653355539
step: 60, loss: 0.00015134115528780967
step: 70, loss: 8.336314931511879e-05
step: 80, loss: 0.00011965718294959515
step: 90, loss: 0.00016202534607145935
step: 100, loss: 0.000733321940060705
step: 110, loss: 8.112486830214038e-05
step: 120, loss: 0.002361917868256569
step: 130, loss: 4.1415969462832436e-05
step: 140, loss: 0.0005329031264409423
step: 150, loss: 0.0019497292814776301
step: 160, loss: 0.0006282167159952223

```

```

step: 170, loss: 0.00014620335423387587
step: 180, loss: 0.0002778259222395718
step: 190, loss: 0.000772689760196954
step: 200, loss: 0.032769475132226944
step: 210, loss: 0.00020724930800497532
step: 220, loss: 0.0029583116993308067
step: 230, loss: 0.000799146480858326
step: 240, loss: 0.00012383359717205167
step: 250, loss: 3.6233424907550216e-05
step: 260, loss: 0.0009759506792761385
step: 270, loss: 0.001459151622839272
step: 280, loss: 0.0002655299613252282
step: 290, loss: 0.0001553100737510249
step: 300, loss: 0.0002943686558865011
step: 310, loss: 0.0016569035360589623
step: 320, loss: 0.0010191358160227537
step: 330, loss: 0.0013135515619069338
step: 340, loss: 0.0033904362935572863
step: 350, loss: 0.00017350148118566722
step: 360, loss: 0.00028385280165821314
step: 370, loss: 0.0012838492402806878
step: 380, loss: 0.005099273752421141
step: 390, loss: 0.00028534213197417557
step: 400, loss: 0.0015744994161650538
step: 410, loss: 0.0016191356116905808
step: 420, loss: 0.00019133204477839172
step: 430, loss: 7.736340921837837e-05
step: 440, loss: 0.012233621440827847
step: 450, loss: 5.876173236174509e-05
step: 460, loss: 8.461297693429515e-05
step: 470, loss: 0.00017706651124171913
step: 480, loss: 8.719958714209497e-05
step: 490, loss: 0.000270206481218338
step: 500, loss: 0.09123064577579498
step: 510, loss: 7.66839730204083e-05
step: 520, loss: 0.00792189035564661
step: 530, loss: 0.004869687370955944
/usr/local/lib/python3.7/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=0.927348449791763, f1=0.9202397418165053,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.022362252697348595
step: 10, loss: 0.0038368890527635813
step: 20, loss: 0.002061122329905629

```

step: 30, loss: 0.024850623682141304
step: 40, loss: 0.0040555899031460285
step: 50, loss: 0.006438758689910173
step: 60, loss: 0.006488774437457323
step: 70, loss: 0.00012699961371254176
step: 80, loss: 0.00015861431893426925
step: 90, loss: 1.9751209038076922e-05
step: 100, loss: 0.0017345501109957695
step: 110, loss: 0.023177146911621094
step: 120, loss: 9.876034164335579e-05
step: 130, loss: 7.953323074616492e-05
step: 140, loss: 0.031296633183956146
step: 150, loss: 6.687342829536647e-05
step: 160, loss: 0.015240604057908058
step: 170, loss: 8.767500548856333e-05
step: 180, loss: 0.003202693536877632
step: 190, loss: 0.003103413851931691
step: 200, loss: 0.012378614395856857
step: 210, loss: 0.02392008900642395
step: 220, loss: 0.0014102841960266232
step: 230, loss: 0.0009753815247677267
step: 240, loss: 0.0006366357556544244
step: 250, loss: 1.952379898284562e-05
step: 260, loss: 6.175140879349783e-05
step: 270, loss: 9.930308442562819e-05
step: 280, loss: 0.0004925283719785511
step: 290, loss: 0.0010286273900419474
step: 300, loss: 0.002404395956546068
step: 310, loss: 0.0010843131458386779
step: 320, loss: 0.0028237425722181797
step: 330, loss: 0.0005313997389748693
step: 340, loss: 2.0533258066279814e-05
step: 350, loss: 0.022559231147170067
step: 360, loss: 0.0001296534901484847
step: 370, loss: 0.0010597355430945754
step: 380, loss: 2.525982381484937e-05
step: 390, loss: 0.00021720162476412952
step: 400, loss: 0.00013073264562990516
step: 410, loss: 0.00038403633516281843
step: 420, loss: 0.00010721445869421586
step: 430, loss: 1.8152988559450023e-05
step: 440, loss: 0.0003123873902950436
step: 450, loss: 0.00021917521371506155
step: 460, loss: 0.000803980918135494
step: 470, loss: 3.458313221926801e-05
step: 480, loss: 0.0007945887628011405
step: 490, loss: 0.00042362138628959656
step: 500, loss: 0.007345553953200579

```

step: 510, loss: 0.005693704355508089
step: 520, loss: 1.2371579941827804e-05
step: 530, loss: 0.00016500714991707355
/usr/local/lib/python3.7/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=0.9304267161410019, f1=0.9228624535315985,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.95828488864936e-05
step: 10, loss: 4.5295979361981153e-05
step: 20, loss: 0.00028487565577961504
step: 30, loss: 0.0005939225666224957
step: 40, loss: 0.0006355839432217181
step: 50, loss: 4.322097811382264e-05
step: 60, loss: 0.0014097323874011636
step: 70, loss: 0.0009007210028357804
step: 80, loss: 5.658773079630919e-05
step: 90, loss: 2.0495255739660934e-05
step: 100, loss: 0.00014069097233004868
step: 110, loss: 0.029881676658988
step: 120, loss: 7.123948307707906e-05
step: 130, loss: 0.005305778235197067
step: 140, loss: 3.13662167172879e-05
step: 150, loss: 3.448593633947894e-05
step: 160, loss: 2.0753161152242683e-05
step: 170, loss: 2.890630094043445e-05
step: 180, loss: 0.00014921659021638334
step: 190, loss: 0.0017687318613752723
step: 200, loss: 0.0004346230416558683
step: 210, loss: 0.004834584891796112
step: 220, loss: 2.0104849681956694e-05
step: 230, loss: 3.016977643710561e-05
step: 240, loss: 0.00031699243118055165
step: 250, loss: 1.9553504898794927e-05
step: 260, loss: 2.2403108232538216e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 270, loss: 4.558507862384431e-05
step: 280, loss: 0.0007451985729858279
step: 290, loss: 0.00027608906384557486
step: 300, loss: 0.03463573753833771
step: 310, loss: 0.00027348779258318245
step: 320, loss: 2.0063427655259147e-05
step: 330, loss: 0.0005982369766570628
step: 340, loss: 3.252369424444623e-05
step: 350, loss: 0.0007248807232826948

```

```

step: 360, loss: 0.004651464521884918
step: 370, loss: 1.963923750736285e-05
step: 380, loss: 9.746908472152427e-05
step: 390, loss: 2.162471537303645e-05
step: 400, loss: 0.0005023559788241982
step: 410, loss: 0.0003949759411625564
step: 420, loss: 0.0038464281242340803
step: 430, loss: 0.0004885190282948315
step: 440, loss: 0.0006218367489054799
step: 450, loss: 8.741407509660348e-05
step: 460, loss: 1.7240281522390433e-05
step: 470, loss: 5.4305943194776773e-05
step: 480, loss: 4.815360807697289e-05
step: 490, loss: 0.00015305842680390924
step: 500, loss: 3.116873267572373e-05
step: 510, loss: 0.0017905129352584481
step: 520, loss: 0.00014529420877806842
step: 530, loss: 3.202432708349079e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9314179796107508, f1=0.9234338747099768,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0010952414013445377
step: 10, loss: 0.0017523972783237696
step: 20, loss: 0.0010293220402672887
step: 30, loss: 0.000547003117389977
step: 40, loss: 2.4202425265684724e-05
step: 50, loss: 0.00026766007067635655
step: 60, loss: 0.00014570688654202968
step: 70, loss: 5.307127139531076e-05
step: 80, loss: 5.6339478760492057e-05
step: 90, loss: 2.0916873836540617e-05
step: 100, loss: 0.00010070428106701002
step: 110, loss: 0.00034246116410940886
step: 120, loss: 0.00037706599687226117
step: 130, loss: 1.9620758394012228e-05
step: 140, loss: 0.00019849237287417054
step: 150, loss: 7.045800884952769e-05
step: 160, loss: 3.159181142109446e-05
step: 170, loss: 0.0014164750464260578
step: 180, loss: 2.5758683477761224e-05
step: 190, loss: 8.81825399119407e-05
step: 200, loss: 2.2391641323338263e-05
step: 210, loss: 0.00010468206892255694

```



```

step: 220, loss: 0.019589092582464218
step: 230, loss: 6.263492105063051e-05
step: 240, loss: 0.00037853349931538105
step: 250, loss: 2.899703940784093e-05
step: 260, loss: 4.189044557278976e-05
step: 270, loss: 1.510210677224677e-05
step: 280, loss: 3.5140878026140854e-05
step: 290, loss: 9.625293023418635e-05
step: 300, loss: 1.9188233636668883e-05
step: 310, loss: 0.0020423568785190582
step: 320, loss: 0.002004310954362154
step: 330, loss: 0.0013622678816318512
step: 340, loss: 0.0004048287228215486
step: 350, loss: 0.06711915880441666
step: 360, loss: 6.169242260511965e-05
step: 370, loss: 2.185108678531833e-05
step: 380, loss: 0.003358916612342
step: 390, loss: 0.0012685713591054082
step: 400, loss: 8.278964378405362e-05
step: 410, loss: 0.0014452970353886485
step: 420, loss: 5.1913917559431866e-05
step: 430, loss: 0.00984988734126091
step: 440, loss: 0.0030406436417251825
step: 450, loss: 0.00014636333798989654
step: 460, loss: 0.03327853977680206
step: 470, loss: 0.00021641667990479618
step: 480, loss: 7.524935062974691e-05
step: 490, loss: 0.129658043384552
step: 500, loss: 0.00017147246398963034
step: 510, loss: 0.00011355571041349322
step: 520, loss: 0.0023754085414111614
step: 530, loss: 0.00018728662689682096
/usr/local/lib/python3.7/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=0.9290382819794585, f1=0.922071861875875,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004557681386359036
step: 10, loss: 3.9286638639168814e-05
step: 20, loss: 0.00016034967848099768
step: 30, loss: 5.466366928885691e-05
step: 40, loss: 0.000247269548708573
step: 50, loss: 0.0009115785942412913
step: 60, loss: 1.3477896573022008e-05
step: 70, loss: 0.00015978410374373198

```

step: 80, loss: 9.087024227483198e-05
step: 90, loss: 8.732245623832569e-05
step: 100, loss: 0.0007566530839540064
step: 110, loss: 3.382021895959042e-05
step: 120, loss: 0.008317249827086926
step: 130, loss: 0.002157095354050398
step: 140, loss: 0.0012873171363025904
step: 150, loss: 6.605660746572539e-05
step: 160, loss: 0.025873471051454544
step: 170, loss: 0.00061472796369344
step: 180, loss: 2.70941382041201e-05
step: 190, loss: 0.0004671672359108925
step: 200, loss: 0.027678649872541428
step: 210, loss: 5.534254523809068e-05
step: 220, loss: 2.2947149773244746e-05
step: 230, loss: 6.192325963638723e-05
step: 240, loss: 0.10745718330144882
step: 250, loss: 3.567552994354628e-05
step: 260, loss: 0.0006414758390747011
step: 270, loss: 0.00012091578537365422
step: 280, loss: 4.832372724195011e-05
step: 290, loss: 0.0014299445319920778
step: 300, loss: 0.000925149885006249
step: 310, loss: 0.018660271540284157
step: 320, loss: 0.0007066850084811449
step: 330, loss: 0.0003909215156454593
step: 340, loss: 0.00045982006122358143
step: 350, loss: 0.0005293504800647497
step: 360, loss: 0.00017719704192131758
step: 370, loss: 8.2128492067568e-05
step: 380, loss: 0.0004041957436129451
step: 390, loss: 0.05272939056158066
step: 400, loss: 0.0006339288665913045
step: 410, loss: 2.530860365368426e-05
step: 420, loss: 0.00012395209341775626
step: 430, loss: 0.03911793604493141
step: 440, loss: 0.00026262577739544213
step: 450, loss: 0.00039630825631320477
step: 460, loss: 0.0007972960011102259
step: 470, loss: 3.4109481930499896e-05
step: 480, loss: 1.5567704394925386e-05
step: 490, loss: 2.2089619960752316e-05
step: 500, loss: 0.002170903142541647
step: 510, loss: 5.6270444474648684e-05
step: 520, loss: 0.00038709226646460593
step: 530, loss: 0.00017303212371189147
/usr/local/lib/python3.7/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=0.931390977443609, f1=0.9233662435354959,
best_f1=0.9259606373008434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021284574177116156
step: 10, loss: 3.6727207771036774e-05
step: 20, loss: 0.00010582920367596671
step: 30, loss: 2.668967135832645e-05
step: 40, loss: 0.09936359524726868
step: 50, loss: 0.0006512815016321838
step: 60, loss: 0.0001636846427572891
step: 70, loss: 0.0026942340191453695
step: 80, loss: 4.2082701838808134e-05
step: 90, loss: 4.00913086195942e-05
step: 100, loss: 6.52499875286594e-05
step: 110, loss: 9.146879892796278e-05
step: 120, loss: 0.0003484625485725701
step: 130, loss: 0.03565428778529167
step: 140, loss: 4.4946136767975986e-05
step: 150, loss: 2.644812411745079e-05
step: 160, loss: 0.00013446535740513355
step: 170, loss: 2.0026613128720783e-05
step: 180, loss: 4.752800305141136e-05
step: 190, loss: 9.08120273379609e-05
step: 200, loss: 0.00010521128569962457
step: 210, loss: 0.001243765465915203
step: 220, loss: 0.005777209531515837
step: 230, loss: 6.990303518250585e-05
step: 240, loss: 5.3358358854893595e-05
step: 250, loss: 6.139335164334625e-05
step: 260, loss: 0.00029409799026325345
step: 270, loss: 1.595510730112437e-05
step: 280, loss: 6.906141788931563e-05
step: 290, loss: 1.46289912663633e-05
step: 300, loss: 6.685619155177847e-05
step: 310, loss: 3.327314698253758e-05
step: 320, loss: 5.4006213758839294e-05
step: 330, loss: 0.00016589846927672625
step: 340, loss: 3.1633364415029064e-05
step: 350, loss: 5.937884998274967e-05
step: 360, loss: 0.001380734029226005
step: 370, loss: 1.2900557521788869e-05
step: 380, loss: 3.1336974643636495e-05
step: 390, loss: 3.51572671206668e-05
step: 400, loss: 4.593687117449008e-05
step: 410, loss: 4.876240564044565e-05
```

```

step: 420, loss: 5.522774517885409e-05
step: 430, loss: 0.0002819288638420403
step: 440, loss: 0.00019617073121480644
step: 450, loss: 1.4852397725917399e-05
step: 460, loss: 1.7180556824314408e-05
step: 470, loss: 0.00123534572776407
step: 480, loss: 4.342197280493565e-05
step: 490, loss: 1.4975435988162644e-05
step: 500, loss: 8.357119077118114e-05
step: 510, loss: 7.274094241438434e-05
step: 520, loss: 0.00028515540179796517
step: 530, loss: 0.0002927251043729484
/usr/local/lib/python3.7/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=0.9311141932501156, f1=0.9249999999999999,
best_f1=0.9259606373008434
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:23, 244.77it/s]

```

```
load_f1 = 0.932963476652797
real_f1 = 0.9308118081180812
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 [00:18, 242.85it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.5591502785682678
step: 10, loss: 0.3613527715206146
step: 20, loss: 0.39411234855651855
step: 30, loss: 0.3216361105442047
step: 40, loss: 0.18724118173122406
step: 50, loss: 0.43697598576545715
step: 60, loss: 0.28674712777137756
step: 70, loss: 0.20105627179145813
step: 80, loss: 0.17675067484378815
step: 90, loss: 0.33595386147499084
step: 100, loss: 0.3378739655017853
step: 110, loss: 0.23076455295085907
step: 120, loss: 0.23276865482330322
step: 130, loss: 0.23462887108325958
step: 140, loss: 0.16104814410209656
```

```

step: 150, loss: 0.23408181965351105
step: 160, loss: 0.23178400099277496
step: 170, loss: 0.31275424361228943
step: 180, loss: 0.16943374276161194
step: 190, loss: 0.18166367709636688
step: 200, loss: 0.27086660265922546
step: 210, loss: 0.24833205342292786
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5514950166112956, f1=0.5912162162162163,
best_f1=0.5912162162162163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1649065911769867
step: 10, loss: 0.24398532509803772
step: 20, loss: 0.18750892579555511
step: 30, loss: 0.1821415275335312
step: 40, loss: 0.2662965655326843
step: 50, loss: 0.18595688045024872
step: 60, loss: 0.3547584116458893
step: 70, loss: 0.11711627244949341
step: 80, loss: 0.23342519998550415
step: 90, loss: 0.13851754367351532
step: 100, loss: 0.018960338085889816
step: 110, loss: 0.13626286387443542
step: 120, loss: 0.12121453136205673
step: 130, loss: 0.047426048666238785
step: 140, loss: 0.1610105186700821
step: 150, loss: 0.17077048122882843
step: 160, loss: 0.24461747705936432
step: 170, loss: 0.11932774633169174
step: 180, loss: 0.22440169751644135
step: 190, loss: 0.22837390005588531
step: 200, loss: 0.03139519318938255
step: 210, loss: 0.18470245599746704
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6083499005964215, f1=0.625250501002004,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045591261237859726
step: 10, loss: 0.1278681606054306
step: 20, loss: 0.19060535728931427
step: 30, loss: 0.1658904254436493
step: 40, loss: 0.09354937821626663
step: 50, loss: 0.036323755979537964
step: 60, loss: 0.24660800397396088
step: 70, loss: 0.13405925035476685
step: 80, loss: 0.26943400502204895
step: 90, loss: 0.04171594977378845
step: 100, loss: 0.15771688520908356
step: 110, loss: 0.12751291692256927
step: 120, loss: 0.1586630493402481
step: 130, loss: 0.2060648500919342
step: 140, loss: 0.10663919150829315
step: 150, loss: 0.19242489337921143
step: 160, loss: 0.05317793786525726
step: 170, loss: 0.21696029603481293
step: 180, loss: 0.13553883135318756
step: 190, loss: 0.16403083503246307
step: 200, loss: 0.07264146208763123
step: 210, loss: 0.15116764605045319
/usr/local/lib/python3.7/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=0.593625498007968, f1=0.6403162055335969,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13267377018928528
step: 10, loss: 0.05697869509458542
step: 20, loss: 0.10906592011451721
step: 30, loss: 0.07549665868282318
step: 40, loss: 0.025530178099870682
step: 50, loss: 0.15391141176223755
step: 60, loss: 0.2652380168437958
step: 70, loss: 0.27657848596572876
step: 80, loss: 0.11108837276697159
step: 90, loss: 0.018197782337665558
step: 100, loss: 0.2990138828754425
step: 110, loss: 0.1427747756242752
step: 120, loss: 0.1481386125087738

```



```

step: 130, loss: 0.14254054427146912
step: 140, loss: 0.2310260385274887
step: 150, loss: 0.07445728033781052
step: 160, loss: 0.06962702423334122
step: 170, loss: 0.20345048606395721
step: 180, loss: 0.41849446296691895
step: 190, loss: 0.0881384089589119
step: 200, loss: 0.0822555273771286
step: 210, loss: 0.20363013446331024
/usr/local/lib/python3.7/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=0.5968379446640317, f1=0.6144578313253013,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11260853707790375
step: 10, loss: 0.06000814214348793
step: 20, loss: 0.2354642003774643
step: 30, loss: 0.06631727516651154
step: 40, loss: 0.04398727789521217
step: 50, loss: 0.041064053773880005
step: 60, loss: 0.07252753525972366
step: 70, loss: 0.10024722665548325
step: 80, loss: 0.08362053334712982
step: 90, loss: 0.053299225866794586
step: 100, loss: 0.002173184184357524
step: 110, loss: 0.13400733470916748
step: 120, loss: 0.12565433979034424
step: 130, loss: 0.09779592603445053
step: 140, loss: 0.20859694480895996
step: 150, loss: 0.04487431421875954
step: 160, loss: 0.23332437872886658
step: 170, loss: 0.18565227091312408
step: 180, loss: 0.1343863308429718
step: 190, loss: 0.012436214834451675
step: 200, loss: 0.12580639123916626
step: 210, loss: 0.01407389435917139
/usr/local/lib/python3.7/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=0.6056074766355141, f1=0.6151012891344384,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05441299453377724

```

```

step: 10, loss: 0.07025355100631714
step: 20, loss: 0.019229399040341377
step: 30, loss: 0.005415529012680054
step: 40, loss: 0.10630866885185242
step: 50, loss: 0.0041344380006194115
step: 60, loss: 0.10304360091686249
step: 70, loss: 0.02603466808795929
step: 80, loss: 0.18398098647594452
step: 90, loss: 0.07671240717172623
step: 100, loss: 0.02032586932182312
step: 110, loss: 0.07279624044895172
step: 120, loss: 0.1069151982665062
step: 130, loss: 0.1314527690410614
step: 140, loss: 0.1219763457775116
step: 150, loss: 0.03562869876623154
step: 160, loss: 0.07956328988075256
step: 170, loss: 0.1914142221212387
step: 180, loss: 0.03437277674674988
step: 190, loss: 0.1316065788269043
step: 200, loss: 0.0761704370379448
step: 210, loss: 0.048204224556684494
/usr/local/lib/python3.7/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=0.5911504424778761, f1=0.5824561403508771,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023162715137004852
step: 10, loss: 0.008523036725819111
step: 20, loss: 0.2017107605934143
step: 30, loss: 0.0317201130092144
step: 40, loss: 0.08575337380170822
step: 50, loss: 0.07452843338251114
step: 60, loss: 0.07109422236680984
step: 70, loss: 0.02816164866089821
step: 80, loss: 0.03218967095017433
step: 90, loss: 0.026388706639409065
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0032414481975138187
step: 110, loss: 0.21451573073863983
step: 120, loss: 0.2267279326915741
step: 130, loss: 0.012716061435639858
step: 140, loss: 0.01395772211253643
step: 150, loss: 0.00949208065867424
step: 160, loss: 0.0908818319439888
step: 170, loss: 0.009706457145512104

```

```

step: 180, loss: 0.1019054502248764
step: 190, loss: 0.05234857276082039
step: 200, loss: 0.03806641697883606
step: 210, loss: 0.032880209386348724
/usr/local/lib/python3.7/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=0.593939393939394, f1=0.5961538461538463,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028903136029839516
step: 10, loss: 0.15500101447105408
step: 20, loss: 0.007914886809885502
step: 30, loss: 0.11936437338590622
step: 40, loss: 0.009027604945003986
step: 50, loss: 0.011880666017532349
step: 60, loss: 0.07361028343439102
step: 70, loss: 0.022706251591444016
step: 80, loss: 0.012476702220737934
step: 90, loss: 0.1122315526008606
step: 100, loss: 0.1473335176706314
step: 110, loss: 0.085179403424263
step: 120, loss: 0.007880846038460732
step: 130, loss: 0.16739127039909363
step: 140, loss: 0.05408293008804321
step: 150, loss: 0.15496322512626648
step: 160, loss: 0.05019364506006241
step: 170, loss: 0.006828028708696365
step: 180, loss: 0.0722016915678978
step: 190, loss: 0.02914159744977951
step: 200, loss: 0.027768932282924652
step: 210, loss: 0.18416157364845276
/usr/local/lib/python3.7/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=0.6071428571428571, f1=0.62109375, best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04386015236377716
step: 10, loss: 0.07102607190608978
step: 20, loss: 0.0029076773207634687
step: 30, loss: 0.004648316185921431
step: 40, loss: 0.07152880728244781
step: 50, loss: 0.018205761909484863
step: 60, loss: 0.1509484201669693

```

```

step: 70, loss: 0.014810883440077305
step: 80, loss: 0.002002421300858259
step: 90, loss: 0.0013932063011452556
step: 100, loss: 0.010582963936030865
step: 110, loss: 0.035534393042325974
step: 120, loss: 0.10395380109548569
step: 130, loss: 0.004434146918356419
step: 140, loss: 0.04375215247273445
step: 150, loss: 0.157255619764328
step: 160, loss: 0.006731214001774788
step: 170, loss: 0.060211341828107834
step: 180, loss: 0.029613690450787544
step: 190, loss: 0.0013018825557082891
step: 200, loss: 0.08462777733802795
step: 210, loss: 0.01687575876712799
/usr/local/lib/python3.7/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=0.5828092243186583, f1=0.6144329896907217,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01573832519352436
step: 10, loss: 0.026536880061030388
step: 20, loss: 0.0023113638162612915
step: 30, loss: 0.052346136420965195
step: 40, loss: 0.028913790360093117
step: 50, loss: 0.015396859496831894
step: 60, loss: 0.09416355192661285
step: 70, loss: 0.031678393483161926
step: 80, loss: 0.028858410194516182
step: 90, loss: 0.05184203013777733
step: 100, loss: 0.15435612201690674
step: 110, loss: 0.03239277005195618
step: 120, loss: 0.01181041169911623
step: 130, loss: 0.012025140225887299
step: 140, loss: 0.0035267272032797337
step: 150, loss: 0.017039678990840912
step: 160, loss: 0.004432699643075466
step: 170, loss: 0.0037692529149353504
step: 180, loss: 0.13799013197422028
step: 190, loss: 0.31894147396087646
step: 200, loss: 0.02181566320359707
step: 210, loss: 0.07398750633001328
/usr/local/lib/python3.7/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=0.5846153846153846, f1=0.5924276169265034,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013474922627210617
step: 10, loss: 0.15390244126319885
step: 20, loss: 0.027818597853183746
step: 30, loss: 0.0027631535194814205
step: 40, loss: 0.06422076374292374
step: 50, loss: 0.014816771261394024
step: 60, loss: 0.02561783976852894
step: 70, loss: 0.012597897090017796
step: 80, loss: 0.06502343714237213
step: 90, loss: 0.053141456097364426
step: 100, loss: 0.05638173222541809
step: 110, loss: 0.08860575407743454
step: 120, loss: 0.16500775516033173
step: 130, loss: 0.005077691283077002
step: 140, loss: 0.02554749883711338
step: 150, loss: 0.005646026227623224
step: 160, loss: 0.01825859025120735
step: 170, loss: 0.03969348967075348
step: 180, loss: 0.007273429539054632
step: 190, loss: 0.024570705369114876
step: 200, loss: 0.0016347152413800359
step: 210, loss: 0.0009984353091567755
/usr/local/lib/python3.7/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=0.6077586206896552, f1=0.6189473684210527,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1200600191950798
step: 10, loss: 0.00418408727273345
step: 20, loss: 0.08801961690187454
step: 30, loss: 0.0017530692275613546
step: 40, loss: 0.11040445417165756
step: 50, loss: 0.0005586952320300043
step: 60, loss: 0.19840098917484283
step: 70, loss: 0.00265975808724761
step: 80, loss: 0.03006741963326931
step: 90, loss: 0.017563750967383385
step: 100, loss: 0.13851819932460785
step: 110, loss: 0.006219794973731041
step: 120, loss: 0.00197496241889894

```

```

step: 130, loss: 0.007950210012495518
step: 140, loss: 0.05000010132789612
step: 150, loss: 0.006016101222485304
step: 160, loss: 0.0013374147238209844
step: 170, loss: 0.0038259057328104973
step: 180, loss: 0.06941510736942291
step: 190, loss: 0.0893123671412468
step: 200, loss: 0.02144745923578739
step: 210, loss: 0.04871445521712303
/usr/local/lib/python3.7/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=0.5818181818181819, f1=0.6217821782178218,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08095782995223999
step: 10, loss: 0.0011152105871587992
step: 20, loss: 0.019505972042679787
step: 30, loss: 0.023687945678830147
step: 40, loss: 0.032832060009241104
step: 50, loss: 0.02420087531208992
step: 60, loss: 0.004106320906430483
step: 70, loss: 0.03025454841554165
step: 80, loss: 0.013238067738711834
step: 90, loss: 0.008049060590565205
step: 100, loss: 0.003416093299165368
step: 110, loss: 0.003128480864688754
step: 120, loss: 0.02598910592496395
step: 130, loss: 0.002298376290127635
step: 140, loss: 0.004281855653971434
step: 150, loss: 0.001670072553679347
step: 160, loss: 0.030875544995069504
step: 170, loss: 0.003285597078502178
step: 180, loss: 0.03814493864774704
step: 190, loss: 0.0009014732204377651
step: 200, loss: 0.0033673730213195086
step: 210, loss: 0.007379299029707909
/usr/local/lib/python3.7/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=0.5795454545454546, f1=0.5932835820895523,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0028381897136569023

```

```

step: 10, loss: 0.025667831301689148
step: 20, loss: 0.0019910119008272886
step: 30, loss: 0.008918197825551033
step: 40, loss: 0.11539403349161148
step: 50, loss: 0.031124746426939964
step: 60, loss: 0.0019605595152825117
step: 70, loss: 0.006556937005370855
step: 80, loss: 0.016773425042629242
step: 90, loss: 0.01768982969224453
step: 100, loss: 0.0063179535791277885
step: 110, loss: 0.03638530895113945
step: 120, loss: 0.08853685110807419
step: 130, loss: 0.0565265491604805
step: 140, loss: 0.034425996243953705
step: 150, loss: 0.03367842733860016
step: 160, loss: 0.003418096574023366
step: 170, loss: 0.003531690454110503
step: 180, loss: 0.03568505123257637
step: 190, loss: 0.012487329542636871
step: 200, loss: 0.0352676659822464
step: 210, loss: 0.12617981433868408
/usr/local/lib/python3.7/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=0.5857740585774057, f1=0.6118143459915611,
best_f1=0.625250501002004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005472994409501553
step: 10, loss: 0.011466771364212036
step: 20, loss: 0.0016832052497193217
step: 30, loss: 0.044325947761535645
step: 40, loss: 0.0036056996323168278
step: 50, loss: 0.0018275425536558032
step: 60, loss: 0.007100405637174845
step: 70, loss: 0.006366051733493805
step: 80, loss: 0.0044759949669241905
step: 90, loss: 0.04776865616440773
step: 100, loss: 0.007201839704066515
step: 110, loss: 0.005461679771542549
step: 120, loss: 0.0018361721886321902
step: 130, loss: 0.004483080469071865
step: 140, loss: 0.0014329658588394523
step: 150, loss: 0.00812985934317112
step: 160, loss: 0.0036259079352021217
step: 170, loss: 0.001514939940534532
step: 180, loss: 0.0009638717165216804

```

```

step: 190, loss: 0.06109052151441574
step: 200, loss: 0.004372614901512861
step: 210, loss: 0.005876230541616678
/usr/local/lib/python3.7/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=0.5848214285714285, f1=0.6035242290748898,
best_f1=0.625250501002004
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 [00:05, 415.56it/s]
load_f1 = 0.6050420168067228
real_f1 = 0.6004319654427647
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 [00:18, 232.33it/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 bert-base-uncased \
--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 bert-base-uncased \
--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: 0.5754364728927612
step: 10, loss: 0.37274765968322754
step: 20, loss: 0.31797242164611816
step: 30, loss: 0.38777491450309753
step: 40, loss: 0.41840001940727234
step: 50, loss: 0.29489102959632874
step: 60, loss: 0.26926136016845703
step: 70, loss: 0.28857454657554626
step: 80, loss: 0.31272536516189575
step: 90, loss: 0.3145526647567749
step: 100, loss: 0.3247881531715393
step: 110, loss: 0.5963802337646484
step: 120, loss: 0.127916619181633
step: 130, loss: 0.11961735039949417
step: 140, loss: 0.03874047473073006
step: 150, loss: 0.18944121897220612
step: 160, loss: 0.1244734525680542
step: 170, loss: 0.23771129548549652
step: 180, loss: 0.044117845594882965
step: 190, loss: 0.19871702790260315
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6582278481012659, f1=0.6601941747572816,
best_f1=0.6601941747572816
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26037633419036865
step: 10, loss: 0.03783087432384491
step: 20, loss: 0.2921459674835205
step: 30, loss: 0.17619511485099792
step: 40, loss: 0.13289429247379303
step: 50, loss: 0.13138295710086823
step: 60, loss: 0.2308504581451416
step: 70, loss: 0.1576625406742096
step: 80, loss: 0.1568375527858734
step: 90, loss: 0.1833440363407135
step: 100, loss: 0.025409035384655
step: 110, loss: 0.19905102252960205
step: 120, loss: 0.2394402027130127
step: 130, loss: 0.04824410006403923
step: 140, loss: 0.07908527553081512
step: 150, loss: 0.09054945409297943
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.025986745953559875
step: 170, loss: 0.1799587458372116
step: 180, loss: 0.2691480815410614
step: 190, loss: 0.04906884580850601
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7079207920792079, f1=0.7281795511221945,
best_f1=0.7281795511221945
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08878526091575623
step: 10, loss: 0.05932341888546944
step: 20, loss: 0.038496989756822586
step: 30, loss: 0.024085767567157745
step: 40, loss: 0.09293122589588165
step: 50, loss: 0.28392305970191956

```

```

step: 60, loss: 0.08300559222698212
step: 70, loss: 0.21386893093585968
step: 80, loss: 0.1338077336549759
step: 90, loss: 0.12979887425899506
step: 100, loss: 0.06947584450244904
step: 110, loss: 0.01303180493414402
step: 120, loss: 0.027393827214837074
step: 130, loss: 0.010716382414102554
step: 140, loss: 0.014396742917597294
step: 150, loss: 0.11635508388280869
step: 160, loss: 0.047177549451589584
step: 170, loss: 0.15652401745319366
step: 180, loss: 0.0773048922419548
step: 190, loss: 0.14031147956848145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7643979057591622, f1=0.7745358090185677,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.048203226178884506
step: 10, loss: 0.1501987874507904
step: 20, loss: 0.24852940440177917
step: 30, loss: 0.03003346547484398
step: 40, loss: 0.03034527786076069
step: 50, loss: 0.05675509572029114
step: 60, loss: 0.03240597993135452
step: 70, loss: 0.33701956272125244
step: 80, loss: 0.33305782079696655
step: 90, loss: 0.03773710876703262
step: 100, loss: 0.06160951778292656
step: 110, loss: 0.03761230409145355
step: 120, loss: 0.05093357712030411
step: 130, loss: 0.17178459465503693
step: 140, loss: 0.0440576933324337
step: 150, loss: 0.12456435710191727
step: 160, loss: 0.032967451959848404
step: 170, loss: 0.03734033927321434
step: 180, loss: 0.008512976579368114
step: 190, loss: 0.3473653495311737
/usr/local/lib/python3.7/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=0.7365728900255755, f1=0.7539267015706806,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.045702770352363586
step: 10, loss: 0.006118652876466513
step: 20, loss: 0.0210807453840971
step: 30, loss: 0.012880942784249783
step: 40, loss: 0.038603682070970535
step: 50, loss: 0.05207731947302818
step: 60, loss: 0.0715387761592865
step: 70, loss: 0.030235914513468742
step: 80, loss: 0.021247612312436104
step: 90, loss: 0.027065740898251534
step: 100, loss: 0.004803547170013189
step: 110, loss: 0.012857906520366669
step: 120, loss: 0.00398228969424963
step: 130, loss: 0.08527609705924988
step: 140, loss: 0.021487023681402206
step: 150, loss: 0.0024724891409277916
step: 160, loss: 0.014997152611613274
step: 170, loss: 0.01653899811208248
step: 180, loss: 0.337453693151474
step: 190, loss: 0.04702224209904671
/usr/local/lib/python3.7/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=0.7540106951871658, f1=0.7444444444444444,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01593375951051712
step: 10, loss: 0.023716069757938385
step: 20, loss: 0.07532109320163727
step: 30, loss: 0.0017495391657575965
step: 40, loss: 0.050738271325826645
step: 50, loss: 0.03626423329114914
step: 60, loss: 0.010376734659075737
step: 70, loss: 0.15959316492080688
step: 80, loss: 0.03377803787589073
step: 90, loss: 0.04377557337284088
step: 100, loss: 0.0013181017711758614
step: 110, loss: 0.017013540491461754
step: 120, loss: 0.12305683642625809
step: 130, loss: 0.015220696106553078

```

```

step: 140, loss: 0.176267609000206
step: 150, loss: 0.0018706442788243294
step: 160, loss: 0.009925637394189835
step: 170, loss: 0.017540384083986282
step: 180, loss: 0.006281902547925711
step: 190, loss: 0.026079688221216202
/usr/local/lib/python3.7/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=0.7365591397849462, f1=0.7252747252747253,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01578487642109394
step: 10, loss: 0.003496079472824931
step: 20, loss: 0.013002662919461727
step: 30, loss: 0.018900863826274872
step: 40, loss: 0.007132804486900568
step: 50, loss: 0.11181433498859406
step: 60, loss: 0.006253461353480816
step: 70, loss: 0.007701423484832048
step: 80, loss: 0.03723960369825363
step: 90, loss: 0.00292415008880198
step: 100, loss: 0.0013104141689836979
step: 110, loss: 0.0042153713293373585
step: 120, loss: 0.002991153858602047
step: 130, loss: 0.001926500117406249
step: 140, loss: 0.01873941719532013
step: 150, loss: 0.03232698142528534
step: 160, loss: 0.030232563614845276
step: 170, loss: 0.004448912106454372
step: 180, loss: 0.0005568084889091551
step: 190, loss: 0.05468330904841423
/usr/local/lib/python3.7/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=0.7539267015706806, f1=0.765498652291105,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.035207733511924744
step: 10, loss: 0.005920293275266886
step: 20, loss: 0.0009519823361188173
step: 30, loss: 0.03827248141169548
step: 40, loss: 0.11022505164146423
step: 50, loss: 0.00623162230476737

```

```

step: 60, loss: 0.07783018052577972
step: 70, loss: 0.003960049711167812
step: 80, loss: 0.0006186144892126322
step: 90, loss: 0.010671843774616718
step: 100, loss: 0.0020144630689173937
step: 110, loss: 0.0010518807685002685
step: 120, loss: 0.00224191602319479
step: 130, loss: 0.002175620524212718
step: 140, loss: 0.011253807693719864
step: 150, loss: 0.041598230600357056
step: 160, loss: 0.0024811681360006332
step: 170, loss: 0.01257340982556343
step: 180, loss: 0.0205396581441164
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 190, loss: 0.004808906465768814
/usr/local/lib/python3.7/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=0.7526881720430108, f1=0.7439353099730459,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002952330047264695
step: 10, loss: 0.21507377922534943
step: 20, loss: 0.008034517988562584
step: 30, loss: 0.002143464284017682
step: 40, loss: 0.003939331509172916
step: 50, loss: 0.005708572920411825
step: 60, loss: 0.0005106689059175551
step: 70, loss: 0.0003350301121827215
step: 80, loss: 0.0011463987175375223
step: 90, loss: 0.003389565274119377
step: 100, loss: 0.06567111611366272
step: 110, loss: 0.0014302728231996298
step: 120, loss: 0.0023578708060085773
step: 130, loss: 0.0306412260979414
step: 140, loss: 0.0069315265864133835
step: 150, loss: 0.0021470419596880674
step: 160, loss: 0.0006488671060651541
step: 170, loss: 0.012831378728151321
step: 180, loss: 0.0598333477973938
step: 190, loss: 0.0007076472975313663
/usr/local/lib/python3.7/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=0.7582417582417582, f1=0.7771739130434783,
best_f1=0.7745358090185677

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0036215006839483976
step: 10, loss: 0.050187818706035614
step: 20, loss: 0.006155471317470074
step: 30, loss: 0.0008961513522081077
step: 40, loss: 0.05966255068778992
step: 50, loss: 0.0023726671934127808
step: 60, loss: 0.15314781665802002
step: 70, loss: 0.01478331908583641
step: 80, loss: 0.0006805305602028966
step: 90, loss: 0.00259325560182333
step: 100, loss: 0.001779771875590086
step: 110, loss: 0.0010314921382814646
step: 120, loss: 0.18224453926086426
step: 130, loss: 0.002317753853276372
step: 140, loss: 0.03699515014886856
step: 150, loss: 0.011227714829146862
step: 160, loss: 0.0009746207506395876
step: 170, loss: 0.0005534274969249964
step: 180, loss: 0.0037149081472307444
step: 190, loss: 0.0010712611256167293
/usr/local/lib/python3.7/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=0.7507002801120448, f1=0.7843137254901961,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002747817779891193
step: 10, loss: 0.0005502650747075677
step: 20, loss: 0.006647662725299597
step: 30, loss: 0.01450225431472063
step: 40, loss: 0.0007770500378683209
step: 50, loss: 0.005278327967971563
step: 60, loss: 0.0007083794916979969
step: 70, loss: 0.0007962769595906138
step: 80, loss: 0.003490337636321783
step: 90, loss: 0.0003993635473307222
step: 100, loss: 0.00044244443415664136
step: 110, loss: 0.00471471669152379
step: 120, loss: 0.0011786588001996279
step: 130, loss: 0.000720759155228734
step: 140, loss: 0.0020894978661090136
step: 150, loss: 0.0007823261548765004
step: 160, loss: 0.0018853646470233798

```



```

step: 170, loss: 0.00054491515038535
step: 180, loss: 0.0004602793778758496
step: 190, loss: 0.0017370289424434304
/usr/local/lib/python3.7/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=0.7548209366391184, f1=0.7821229050279329,
best_f1=0.7745358090185677
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000996092800050974
step: 10, loss: 0.005592815112322569
step: 20, loss: 0.0007136081112548709
step: 30, loss: 0.002792682498693466
step: 40, loss: 0.015257886610925198
step: 50, loss: 0.002441779710352421
step: 60, loss: 0.02063959650695324
step: 70, loss: 0.004388793371617794
step: 80, loss: 0.003128783078864217
step: 90, loss: 0.0027037356048822403
step: 100, loss: 0.0007241209968924522
step: 110, loss: 0.0007264756131917238
step: 120, loss: 0.0006471549277193844
step: 130, loss: 0.003468045499175787
step: 140, loss: 0.0033049657940864563
step: 150, loss: 0.00045071454951539636
step: 160, loss: 0.012760859914124012
step: 170, loss: 0.004614532925188541
step: 180, loss: 0.0004950602306053042
step: 190, loss: 0.0010553475003689528
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7717391304347826, f1=0.7683923705722071,
best_f1=0.7683923705722071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013539273058995605
step: 10, loss: 0.008541020564734936
step: 20, loss: 0.0004981959355063736
step: 30, loss: 0.008011904545128345
step: 40, loss: 0.0005605563637800515

```

```

step: 50, loss: 0.0022754520177841187
step: 60, loss: 0.0006143946666270494
step: 70, loss: 0.040749140083789825
step: 80, loss: 0.008408461697399616
step: 90, loss: 0.005914660636335611
step: 100, loss: 0.026829397305846214
step: 110, loss: 0.0009821048006415367
step: 120, loss: 0.0013852481497451663
step: 130, loss: 0.0025408139917999506
step: 140, loss: 0.0004097140918020159
step: 150, loss: 0.0008089351467788219
step: 160, loss: 0.0009586484520696104
step: 170, loss: 0.0005699563771486282
step: 180, loss: 0.0012870568316429853
step: 190, loss: 0.001294466550461948
/usr/local/lib/python3.7/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=0.7692307692307692, f1=0.7777777777777778,
best_f1=0.7683923705722071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02436958998441696
step: 10, loss: 0.0012322379043325782
step: 20, loss: 0.0007447054958902299
step: 30, loss: 0.00042913787183351815
step: 40, loss: 0.0030688161496073008
step: 50, loss: 0.0005830790614709258
step: 60, loss: 0.0007623675046488643
step: 70, loss: 0.0009539610473439097
step: 80, loss: 0.008611004799604416
step: 90, loss: 0.0022218339145183563
step: 100, loss: 0.0034733039792627096
step: 110, loss: 0.004033922683447599
step: 120, loss: 0.1413804292678833
step: 130, loss: 0.00281276972964406
step: 140, loss: 0.10581211000680923
step: 150, loss: 0.0010144587140530348
step: 160, loss: 0.0038661437574774027
step: 170, loss: 0.0012007084442302585
step: 180, loss: 0.005690384656190872
step: 190, loss: 0.0008638322469778359
/usr/local/lib/python3.7/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=0.7594936708860759, f1=0.7512437810945273,
best_f1=0.7683923705722071

```

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

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

```
step: 0, loss: 0.0013674176298081875
step: 10, loss: 0.0005550537607632577
step: 20, loss: 0.0004908666596747935
step: 30, loss: 0.0008842690149322152
step: 40, loss: 0.00031305180164054036
step: 50, loss: 0.003786759916692972
step: 60, loss: 0.008723122999072075
step: 70, loss: 0.00046964408829808235
step: 80, loss: 0.004242732655256987
step: 90, loss: 0.0006857895641587675
step: 100, loss: 0.0116479666903615
step: 110, loss: 0.0010926782852038741
step: 120, loss: 0.00037967541720718145
step: 130, loss: 0.004531232174485922
step: 140, loss: 0.07506348937749863
step: 150, loss: 0.000262886518612504
step: 160, loss: 0.004413124639540911
step: 170, loss: 0.0012763881823047996
step: 180, loss: 0.003920495975762606
step: 190, loss: 0.0008956511155702174
```

```
/usr/local/lib/python3.7/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=0.7563025210084034, f1=0.7627118644067796,
best_f1=0.7683923705722071
```

```
mv: cannot stat '*_dev.pt': No such file or directory
```

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

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

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

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

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

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

```

loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:09, 210.47it/s]
load_f1 = 0.6855670103092784
real_f1 = 0.6616541353383459
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 [00:18, 236.31it/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 bert-base-uncased \
    --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 bert-base-uncased \
--use_gpu \
--fp16 \
--checkpoint_path checkpoints/
```

```
Downloading: 100% 433/433 [00:00<00:00, 594kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 25.1MB/s]
Downloading: 100% 440M/440M [00:06<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.")
```

```
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: 0.6382588148117065
step: 10, loss: 0.3735152781009674
step: 20, loss: 0.3016364574432373
step: 30, loss: 0.39490753412246704
step: 40, loss: 0.2798600196838379
step: 50, loss: 0.25113895535469055
step: 60, loss: 0.23774713277816772
step: 70, loss: 0.3686505854129791
step: 80, loss: 0.3841409385204315
step: 90, loss: 0.2571116089820862
step: 100, loss: 0.27272000908851624
step: 110, loss: 0.21391870081424713
step: 120, loss: 0.3739422857761383
step: 130, loss: 0.06271576136350632
step: 140, loss: 0.1982114315032959
step: 150, loss: 0.24036845564842224
step: 160, loss: 0.14275866746902466
step: 170, loss: 0.18395505845546722
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6603325415676958, f1=0.6588785046728973,
best_f1=0.6588785046728973
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20569409430027008
step: 10, loss: 0.18306735157966614
step: 20, loss: 0.12092436105012894
step: 30, loss: 0.1690874546766281
step: 40, loss: 0.04890768229961395
step: 50, loss: 0.10741548985242844
step: 60, loss: 0.11618684977293015
step: 70, loss: 0.08915266394615173
step: 80, loss: 0.04381820932030678
step: 90, loss: 0.0768558457493782
step: 100, loss: 0.25695720314979553
step: 110, loss: 0.06408365070819855
step: 120, loss: 0.10167885571718216
step: 130, loss: 0.1486247330904007
step: 140, loss: 0.30544617772102356
step: 150, loss: 0.1608988493680954

```

```

step: 160, loss: 0.19830051064491272
step: 170, loss: 0.07241767644882202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7643979057591622, f1=0.7543424317617866,
best_f1=0.7543424317617866
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1771627813577652
step: 10, loss: 0.2019398957490921
step: 20, loss: 0.06437958776950836
step: 30, loss: 0.21090947091579437
step: 40, loss: 0.10232047736644745
step: 50, loss: 0.10335147380828857
step: 60, loss: 0.21263235807418823
step: 70, loss: 0.06719411164522171
step: 80, loss: 0.03580252453684807
step: 90, loss: 0.2759377956390381
step: 100, loss: 0.05085710808634758
step: 110, loss: 0.045040614902973175
step: 120, loss: 0.1273248940706253
step: 130, loss: 0.19810569286346436
step: 140, loss: 0.1661810576915741
step: 150, loss: 0.011889401823282242
step: 160, loss: 0.00785076804459095
step: 170, loss: 0.09051244705915451
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7832512315270935, f1=0.772093023255814,
best_f1=0.772093023255814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010548298247158527
step: 10, loss: 0.03318019211292267
step: 20, loss: 0.025857439264655113
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```

```

step: 30, loss: 0.02258303388953209
step: 40, loss: 0.008005013689398766
step: 50, loss: 0.04870303347706795
step: 60, loss: 0.24029922485351562
step: 70, loss: 0.01448972336947918
step: 80, loss: 0.04867205396294594
step: 90, loss: 0.1414155662059784
step: 100, loss: 0.22290083765983582
step: 110, loss: 0.274906724691391
step: 120, loss: 0.008361630141735077
step: 130, loss: 0.05368444696068764
step: 140, loss: 0.07393131405115128
step: 150, loss: 0.19539028406143188
step: 160, loss: 0.054271988570690155
step: 170, loss: 0.006403012666851282
/usr/local/lib/python3.7/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=0.7780548628428928, f1=0.7864077669902912,
best_f1=0.772093023255814
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03532961755990982
step: 10, loss: 0.016211099922657013
step: 20, loss: 0.06329654902219772
step: 30, loss: 0.031524017453193665
step: 40, loss: 0.027151847258210182
step: 50, loss: 0.16494455933570862
step: 60, loss: 0.08793344348669052
step: 70, loss: 0.1436934769153595
step: 80, loss: 0.25962185859680176
step: 90, loss: 0.07049696147441864
step: 100, loss: 0.2611754834651947
step: 110, loss: 0.1556156426668167
step: 120, loss: 0.05373804271221161
step: 130, loss: 0.04895049333572388
step: 140, loss: 0.00290064699947834
step: 150, loss: 0.032980777323246
step: 160, loss: 0.02752435952425003
step: 170, loss: 0.01210855320096016
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8166259168704155, f1=0.8076009501187648,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016344113275408745
step: 10, loss: 0.021594736725091934
step: 20, loss: 0.03043089434504509
step: 30, loss: 0.022987335920333862
step: 40, loss: 0.007218166720122099
step: 50, loss: 0.2372358739376068
step: 60, loss: 0.1054198145866394
step: 70, loss: 0.18157272040843964
step: 80, loss: 0.056406378746032715
step: 90, loss: 0.038307882845401764
step: 100, loss: 0.01019169669598341
step: 110, loss: 0.0014121334534138441
step: 120, loss: 0.03763160482048988
step: 130, loss: 0.036635588854551315
step: 140, loss: 0.12697049975395203
step: 150, loss: 0.024159051477909088
step: 160, loss: 0.12716637551784515
step: 170, loss: 0.07398227602243423
/usr/local/lib/python3.7/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=0.7830687830687831, f1=0.761904761904762,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02028767764568329
step: 10, loss: 0.005131332203745842
step: 20, loss: 0.07159207761287689
step: 30, loss: 0.004372369032353163
step: 40, loss: 0.024927766993641853
step: 50, loss: 0.033300913870334625
step: 60, loss: 0.006300886627286673
step: 70, loss: 0.002125043887645006
step: 80, loss: 0.01694617047905922
step: 90, loss: 0.00323398201726377
step: 100, loss: 0.0005338441696949303
step: 110, loss: 0.19125881791114807
step: 120, loss: 0.05573524534702301
step: 130, loss: 0.10902043431997299
step: 140, loss: 0.18522877991199493
step: 150, loss: 0.009846481494605541
step: 160, loss: 0.08686904609203339

```

```

step: 170, loss: 0.16354775428771973
/usr/local/lib/python3.7/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=0.7941176470588236, f1=0.7713625866050808,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011609701439738274
step: 10, loss: 0.028374915942549706
step: 20, loss: 0.0066437809728085995
step: 30, loss: 0.0238620787858963
step: 40, loss: 0.0003630045976024121
step: 50, loss: 0.005208578426390886
step: 60, loss: 0.005066650453954935
step: 70, loss: 0.03749929368495941
step: 80, loss: 0.01736893132328987
step: 90, loss: 0.012864232063293457
step: 100, loss: 0.03616498410701752
step: 110, loss: 0.07423548400402069
step: 120, loss: 0.00868464820086956
step: 130, loss: 0.004862077534198761
step: 140, loss: 0.007677697576582432
step: 150, loss: 0.30763518810272217
step: 160, loss: 0.005310962442308664
step: 170, loss: 0.0493488535284996
/usr/local/lib/python3.7/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=0.7858942065491185, f1=0.7677725118483412,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004167004663031548
step: 10, loss: 0.08652154356241226
step: 20, loss: 0.004325177054852247
step: 30, loss: 0.0270827803760767
step: 40, loss: 0.003389936638996005
step: 50, loss: 0.0006410720525309443
step: 60, loss: 0.08585745841264725
step: 70, loss: 0.014046336524188519
step: 80, loss: 0.0009947381913661957
step: 90, loss: 0.025737429037690163
step: 100, loss: 0.06945355981588364
step: 110, loss: 0.007418195251375437
step: 120, loss: 0.0031481250189244747

```

```

step: 130, loss: 0.03264731168746948
step: 140, loss: 0.13159678876399994
step: 150, loss: 0.0028419073205441236
step: 160, loss: 0.1538509875535965
step: 170, loss: 0.021081287413835526
/usr/local/lib/python3.7/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=0.7980295566502462, f1=0.7741935483870966,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07351604849100113
step: 10, loss: 0.002565047238022089
step: 20, loss: 0.07928292453289032
step: 30, loss: 0.012552985921502113
step: 40, loss: 0.0002325509412912652
step: 50, loss: 0.047606877982616425
step: 60, loss: 0.0003409194468986243
step: 70, loss: 0.0016318848356604576
step: 80, loss: 0.0002706522645894438
step: 90, loss: 0.009161020629107952
step: 100, loss: 0.0010724577587097883
step: 110, loss: 0.0008367894333787262
step: 120, loss: 0.023019710555672646
step: 130, loss: 0.0416899211704731
step: 140, loss: 0.0026076447684317827
step: 150, loss: 0.019716234877705574
step: 160, loss: 0.009303918108344078
step: 170, loss: 0.001285894657485187
/usr/local/lib/python3.7/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=0.7918781725888325, f1=0.7772511848341233,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02750048041343689
step: 10, loss: 0.006612843368202448
step: 20, loss: 0.0014381235232576728
step: 30, loss: 0.0004373996052891016
step: 40, loss: 0.0007802999462001026
step: 50, loss: 0.004636287689208984
step: 60, loss: 0.06557587534189224
step: 70, loss: 0.00039361976087093353
step: 80, loss: 0.0006167241372168064

```

```

step: 90, loss: 0.002802010625600815
step: 100, loss: 0.0010802808683365583
step: 110, loss: 0.027742763981223106
step: 120, loss: 0.0006091149407438934
step: 130, loss: 0.0002473840431775898
step: 140, loss: 0.00316755217500031
step: 150, loss: 0.007609803695231676
step: 160, loss: 0.026591569185256958
step: 170, loss: 0.0011070745531469584
/usr/local/lib/python3.7/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=0.8095238095238095, f1=0.7555555555555555,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015222594141960144
step: 10, loss: 0.000613891170360148
step: 20, loss: 0.00042807511636056006
step: 30, loss: 0.0019926163367927074
step: 40, loss: 0.0027604647912085056
step: 50, loss: 0.0027726374100893736
step: 60, loss: 0.014198647812008858
step: 70, loss: 0.06885785609483719
step: 80, loss: 0.0007278930279426277
step: 90, loss: 0.0027084939647465944
step: 100, loss: 0.002341283019632101
step: 110, loss: 0.0026189018972218037
step: 120, loss: 0.0010365914786234498
step: 130, loss: 0.06958725303411484
step: 140, loss: 0.007576644420623779
step: 150, loss: 0.005858494900166988
step: 160, loss: 0.0007107358542270958
step: 170, loss: 0.0004820397007279098
/usr/local/lib/python3.7/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=0.762402088772846, f1=0.7853658536585365,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23767608404159546
step: 10, loss: 0.05700230970978737
step: 20, loss: 0.003402977716177702
step: 30, loss: 0.0008768919506110251
step: 40, loss: 0.026398399844765663

```

```

step: 50, loss: 0.0012653230223804712
step: 60, loss: 0.009306834079325199
step: 70, loss: 0.020082546398043633
step: 80, loss: 0.004694918170571327
step: 90, loss: 0.000393409573007375
step: 100, loss: 0.006348202005028725
step: 110, loss: 0.000408947526011616
step: 120, loss: 0.027860363945364952
step: 130, loss: 0.0011033340124413371
step: 140, loss: 0.0005995667888782918
step: 150, loss: 0.0019223574781790376
step: 160, loss: 0.10570166260004044
step: 170, loss: 0.0014787674881517887
/usr/local/lib/python3.7/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=0.7830687830687831, f1=0.8061224489795918,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03312757611274719
step: 10, loss: 0.0004603721317835152
step: 20, loss: 0.00744609534740448
step: 30, loss: 0.006881830282509327
step: 40, loss: 0.0006749862222932279
step: 50, loss: 0.07282920926809311
step: 60, loss: 0.002896952209994197
step: 70, loss: 0.0021101871971040964
step: 80, loss: 0.0063686855137348175
step: 90, loss: 0.001964298076927662
step: 100, loss: 0.0023893583565950394
step: 110, loss: 0.10865268111228943
step: 120, loss: 0.055931005626916885
step: 130, loss: 0.04228787124156952
step: 140, loss: 0.004499190021306276
step: 150, loss: 0.0027054203674197197
step: 160, loss: 0.0018933992832899094
step: 170, loss: 0.003037984250113368
/usr/local/lib/python3.7/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=0.7990314769975787,
best_f1=0.8076009501187648
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003691143065225333

```

```

step: 10, loss: 0.00466390373185277
step: 20, loss: 0.04183083772659302
step: 30, loss: 0.0011103228898718953
step: 40, loss: 0.001974657643586397
step: 50, loss: 0.004648052155971527
step: 60, loss: 0.018611283972859383
step: 70, loss: 0.00034575039171613753
step: 80, loss: 0.02585310861468315
step: 90, loss: 0.0007342136232182384
step: 100, loss: 0.007424386218190193
step: 110, loss: 0.000858645944390446
step: 120, loss: 0.0045915404334664345
step: 130, loss: 0.004480821546167135
step: 140, loss: 0.024431873112916946
step: 150, loss: 0.004447432234883308
step: 160, loss: 0.012131715193390846
step: 170, loss: 0.004003035370260477
/usr/local/lib/python3.7/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=0.8110831234256927, f1=0.7971014492753623,
best_f1=0.8076009501187648
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
1916it [00:06, 289.70it/s]
load_f1 = 0.33093525179856115
real_f1 = 0.3290845886442642
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 264.90it/s]

```

4.3 DITTO DIRTY

4.3.1 DBLP-ACM - Running the matcher

```

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

```

```
--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: 0.6036078929901123

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

step: 10, loss: 0.6152273416519165

step: 20, loss: 0.45180797576904297

step: 30, loss: 0.28557056188583374

step: 40, loss: 0.20734496414661407

step: 50, loss: 0.0689634159207344


```

step: 60, loss: 0.1445380300283432
step: 70, loss: 0.09373138844966888
step: 80, loss: 0.09237721562385559
step: 90, loss: 0.16688038408756256
step: 100, loss: 0.01571529358625412
step: 110, loss: 0.21531370282173157
step: 120, loss: 0.014417739585042
step: 130, loss: 0.011900393292307854
step: 140, loss: 0.002966728061437607
step: 150, loss: 0.010401282459497452
step: 160, loss: 0.02881181612610817
step: 170, loss: 0.21406282484531403
step: 180, loss: 0.1759716123342514
step: 190, loss: 0.08901552855968475
step: 200, loss: 0.1024487167596817
step: 210, loss: 0.0043702335096895695
step: 220, loss: 0.024508029222488403
step: 230, loss: 0.05183238536119461
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9682539682539683, f1=0.9608294930875576,
best_f1=0.9608294930875576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011756865307688713
step: 10, loss: 0.010090998373925686
step: 20, loss: 0.11384432762861252
step: 30, loss: 0.19167260825634003
step: 40, loss: 0.03965531289577484
step: 50, loss: 0.004085137043148279
step: 60, loss: 0.018265375867486
step: 70, loss: 0.11246372014284134
step: 80, loss: 0.09244529902935028
step: 90, loss: 0.025097284466028214
step: 100, loss: 0.11751779168844223
step: 110, loss: 0.13281042873859406
step: 120, loss: 0.14326918125152588
step: 130, loss: 0.028123434633016586
step: 140, loss: 0.0035040637012571096
step: 150, loss: 0.012316172942519188
step: 160, loss: 0.017388876527547836
step: 170, loss: 0.002487249206751585

```

```

step: 180, loss: 0.0833701565861702
step: 190, loss: 0.003050851169973612
step: 200, loss: 0.004548192024230957
step: 210, loss: 0.033368874341249466
step: 220, loss: 0.03578527644276619
step: 230, loss: 0.029627203941345215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9765886287625419, f1=0.9743589743589743,
best_f1=0.9743589743589743
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00814746879041195
step: 10, loss: 0.012469637207686901
step: 20, loss: 0.003033985150977969
step: 30, loss: 0.1544475257396698
step: 40, loss: 0.16189265251159668
step: 50, loss: 0.009965581819415092
step: 60, loss: 0.007917391136288643
step: 70, loss: 0.033881817013025284
step: 80, loss: 0.0011874452466145158
step: 90, loss: 0.005773146171122789
step: 100, loss: 0.004999507684260607
step: 110, loss: 0.0010267851175740361
step: 120, loss: 0.055479198694229126
step: 130, loss: 0.0026390256825834513
step: 140, loss: 0.005984729155898094
step: 150, loss: 0.008014963939785957
step: 160, loss: 0.01696452870965004
step: 170, loss: 0.013586190529167652
step: 180, loss: 0.05645343288779259
step: 190, loss: 0.007047305814921856
step: 200, loss: 0.004471295513212681
step: 210, loss: 0.00184512825217098
step: 220, loss: 0.04703930392861366
step: 230, loss: 0.0006231777952052653
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.978912319644839, f1=0.9721913236929923,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001032689237035811
step: 10, loss: 0.0007902246434241533
step: 20, loss: 0.019731566309928894
step: 30, loss: 0.000880143721587956
step: 40, loss: 0.005762617569416761
step: 50, loss: 0.001176354824565351
step: 60, loss: 0.0015313575277104974
step: 70, loss: 0.0009268406429328024
step: 80, loss: 0.0022653997875750065
step: 90, loss: 0.0032435439061373472
step: 100, loss: 0.001043914700858295
step: 110, loss: 0.007316647097468376
step: 120, loss: 0.05768914520740509
step: 130, loss: 0.0017346504610031843
step: 140, loss: 0.018132707104086876
step: 150, loss: 0.14891138672828674
step: 160, loss: 0.11006245017051697
step: 170, loss: 0.006928429938852787
step: 180, loss: 0.0007607451407238841
step: 190, loss: 0.00889984518289566
step: 200, loss: 0.0009509255760349333
step: 210, loss: 0.13595469295978546
step: 220, loss: 0.00035656520049087703
step: 230, loss: 0.006300784647464752
/usr/local/lib/python3.7/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=0.9775784753363228, f1=0.9786276715410572,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007737474516034126
step: 10, loss: 0.024852409958839417
step: 20, loss: 0.002764445962384343
step: 30, loss: 0.0003845811006613076
step: 40, loss: 0.0004556264029815793
step: 50, loss: 0.000492453807964921
step: 60, loss: 0.16097892820835114
step: 70, loss: 0.009141388349235058
step: 80, loss: 0.00048485337174497545
step: 90, loss: 0.008931964635848999

```

```

step: 100, loss: 0.00045525195309892297
step: 110, loss: 0.0010581298265606165
step: 120, loss: 0.0023466141428798437
step: 130, loss: 0.0016209472669288516
step: 140, loss: 0.0035701259039342403
step: 150, loss: 0.0008669144008308649
step: 160, loss: 0.0011627546045929193
step: 170, loss: 0.01963425613939762
step: 180, loss: 0.022440314292907715
step: 190, loss: 0.12412431836128235
step: 200, loss: 0.018044373020529747
step: 210, loss: 0.0050128246657550335
step: 220, loss: 0.0031435585115104914
step: 230, loss: 0.0034496800508350134
/usr/local/lib/python3.7/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=0.9777777777777777, f1=0.9765886287625419,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024016056209802628
step: 10, loss: 0.00027635128935799
step: 20, loss: 0.0006065406487323344
step: 30, loss: 0.00019112932204734534
step: 40, loss: 0.0010742275044322014
step: 50, loss: 0.0005832203896716237
step: 60, loss: 0.006855517625808716
step: 70, loss: 0.0040524061769247055
step: 80, loss: 0.000738044735044241
step: 90, loss: 0.001861167955212295
step: 100, loss: 0.05479247495532036
step: 110, loss: 0.0038012610748410225
step: 120, loss: 0.000976684968918562
step: 130, loss: 0.00025349666248075664
step: 140, loss: 0.0009988733800128102
step: 150, loss: 0.04685477539896965
step: 160, loss: 0.00039396705687977374
step: 170, loss: 0.00030082507873885334
step: 180, loss: 0.014672794379293919
step: 190, loss: 0.0037771398201584816
step: 200, loss: 0.0001003577999654226
step: 210, loss: 0.0006859177374280989
step: 220, loss: 0.00010901251516770571
step: 230, loss: 0.023104393854737282
/usr/local/lib/python3.7/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=0.9755011135857461, f1=0.9765886287625419,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00028315564850345254
step: 10, loss: 0.00011850603186758235
step: 20, loss: 0.0003445638867560774
step: 30, loss: 0.00013752082304563373
step: 40, loss: 0.001010993029922247
step: 50, loss: 0.0003810396883636713
step: 60, loss: 0.009294604882597923
step: 70, loss: 0.0060250586830079556
step: 80, loss: 0.0004123076214455068
step: 90, loss: 0.0003191581345163286
step: 100, loss: 0.0001292716187890619
step: 110, loss: 0.0008423259132541716
step: 120, loss: 0.00027511161169968545
step: 130, loss: 0.000244701310293749
step: 140, loss: 0.00043041701428592205
step: 150, loss: 0.000681597157381475
step: 160, loss: 0.09402839839458466
step: 170, loss: 0.010603900998830795
step: 180, loss: 0.0008323994698002934
step: 190, loss: 0.0005324704106897116
step: 200, loss: 0.04756081476807594
step: 210, loss: 8.444728155154735e-05
step: 220, loss: 0.0003569523396436125
step: 230, loss: 0.00039907643804326653
/usr/local/lib/python3.7/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=0.9765886287625419, f1=0.9776286353467561,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.217625195859e-05
step: 10, loss: 0.0051375748589634895
step: 20, loss: 0.014629172161221504
step: 30, loss: 7.40487957955338e-05
step: 40, loss: 0.006328292191028595
step: 50, loss: 0.0002123154845321551
step: 60, loss: 6.766864680685103e-05
step: 70, loss: 0.0026076713111251593
step: 80, loss: 9.468349890084937e-05
step: 90, loss: 3.443565219640732e-05

```

```

step: 100, loss: 0.0001571990578668192
step: 110, loss: 0.031099315732717514
step: 120, loss: 0.02181997336447239
step: 130, loss: 0.008176887407898903
step: 140, loss: 0.00010288010525982827
step: 150, loss: 7.061901851557195e-05
step: 160, loss: 0.0002481411211192608
step: 170, loss: 0.00013089798449072987
step: 180, loss: 0.0007977747591212392
step: 190, loss: 0.004597879014909267
step: 200, loss: 0.003407923271879554
step: 210, loss: 0.002522081835195422
step: 220, loss: 0.0004047861439175904
step: 230, loss: 0.0002851660829037428
/usr/local/lib/python3.7/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=0.9773242630385486, f1=0.9748858447488584,
best_f1=0.9721913236929923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000531586934812367
step: 10, loss: 0.0013185156276449561
step: 20, loss: 0.00039960749563761055
step: 30, loss: 0.00019168273138348013
step: 40, loss: 0.004598418716341257
step: 50, loss: 0.0003616784233599901
step: 60, loss: 0.0001669190387474373
step: 70, loss: 0.001031421241350472
step: 80, loss: 7.694278610870242e-05
step: 90, loss: 0.0012806139420717955
step: 100, loss: 0.00021733860194217414
step: 110, loss: 5.8438843552721664e-05
step: 120, loss: 6.0039456002414227e-05
step: 130, loss: 0.00014221062883734703
step: 140, loss: 5.088116813567467e-05
step: 150, loss: 0.00018736865604296327
step: 160, loss: 5.30338948010467e-05
step: 170, loss: 7.193767669377849e-05
step: 180, loss: 0.00025283879949711263
step: 190, loss: 0.00011649663065327331
step: 200, loss: 0.0009690753649920225
step: 210, loss: 0.0002960807178169489
step: 220, loss: 0.00013008325186092407
step: 230, loss: 6.907113856868818e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9794988610478361, f1=0.9772727272727272,
best_f1=0.9772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.28502922761254e-05
step: 10, loss: 8.723825158085674e-05
step: 20, loss: 0.0005236138240434229
step: 30, loss: 8.190791413653642e-05
step: 40, loss: 0.0061009894125163555
step: 50, loss: 0.009475363418459892
step: 60, loss: 7.228054164443165e-05
step: 70, loss: 0.002020786749199033
step: 80, loss: 0.000144995326991193
step: 90, loss: 0.0009590164409019053
step: 100, loss: 0.00023019526270218194
step: 110, loss: 0.001922071911394596
step: 120, loss: 9.519093873677775e-05
step: 130, loss: 0.0002002950495807454
step: 140, loss: 0.06417283415794373
step: 150, loss: 0.02628319337964058
step: 160, loss: 0.00021991250105202198
step: 170, loss: 8.019653614610434e-05
step: 180, loss: 0.000388311076676473
step: 190, loss: 0.006344526540488005
step: 200, loss: 8.451961184618995e-05
step: 210, loss: 0.0010915591847151518
step: 220, loss: 0.0003859613789245486
step: 230, loss: 0.0011279081227257848
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9785310734463276, f1=0.9785310734463276,
best_f1=0.9772727272727272
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003414472157601267
step: 10, loss: 0.0001898606715258211
step: 20, loss: 0.0003876727423630655
step: 30, loss: 0.0016443085623905063
step: 40, loss: 0.00034929902176372707
step: 50, loss: 0.00011326761159580201

```

```

step: 60, loss: 0.002068394096568227
step: 70, loss: 0.00022603616525884718
step: 80, loss: 5.873217014595866e-05
step: 90, loss: 0.00011113355139968917
step: 100, loss: 0.0027018056716769934
step: 110, loss: 0.0015743756666779518
step: 120, loss: 6.721351383021101e-05
step: 130, loss: 8.511533815180883e-05
step: 140, loss: 4.594562778947875e-05
step: 150, loss: 0.0006710501038469374
step: 160, loss: 6.374645454343408e-05
step: 170, loss: 0.0699312761425972
step: 180, loss: 0.0007551810122095048
step: 190, loss: 5.767084803665057e-05
step: 200, loss: 0.00015057400742080063
step: 210, loss: 4.461204298422672e-05
step: 220, loss: 6.776185909984633e-05
step: 230, loss: 0.00011119216651422903
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9807037457434733, f1=0.9761634506242906,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.791964551666752e-05
step: 10, loss: 0.00033154728589579463
step: 20, loss: 0.00045829766895622015
step: 30, loss: 0.0003994816215708852
step: 40, loss: 0.00013552462041843683
step: 50, loss: 0.0016194666968658566
step: 60, loss: 0.0005457103834487498
step: 70, loss: 4.7770761739229783e-05
step: 80, loss: 0.0017115080263465643
step: 90, loss: 5.536648677662015e-05
step: 100, loss: 0.00014546074089594185
step: 110, loss: 0.00010550850129220635
step: 120, loss: 0.00021133493282832205
step: 130, loss: 8.144097228068858e-05
step: 140, loss: 5.589555439655669e-05
step: 150, loss: 0.0001614154316484928
step: 160, loss: 5.657095607602969e-05
step: 170, loss: 9.439831046620384e-05

```



```

step: 180, loss: 0.00010083200322696939
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.00013000410399399698
step: 200, loss: 3.954590283683501e-05
step: 210, loss: 5.631167732644826e-05
step: 220, loss: 3.753856071853079e-05
step: 230, loss: 0.026254555210471153
/usr/local/lib/python3.7/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=0.9753363228699552, f1=0.9798657718120806,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011874268238898367
step: 10, loss: 8.899473323253915e-05
step: 20, loss: 0.0001426788221579045
step: 30, loss: 0.00012454397801775485
step: 40, loss: 0.00017640288569964468
step: 50, loss: 7.956816989462823e-05
step: 60, loss: 0.0006769172614440322
step: 70, loss: 4.3111504055559635e-05
step: 80, loss: 4.1233102820115164e-05
step: 90, loss: 0.00013422466872725636
step: 100, loss: 6.442348239943385e-05
step: 110, loss: 0.016497278586030006
step: 120, loss: 0.0016634105704724789
step: 130, loss: 5.926973608438857e-05
step: 140, loss: 0.0001027822945616208
step: 150, loss: 4.26746737502981e-05
step: 160, loss: 5.665925345965661e-05
step: 170, loss: 0.0001047076511895284
step: 180, loss: 0.002264486625790596
step: 190, loss: 4.07858460675925e-05
step: 200, loss: 0.0005867459112778306
step: 210, loss: 3.3875301596708596e-05
step: 220, loss: 2.9336029911064543e-05
step: 230, loss: 4.906893809675239e-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=0.9785310734463276, f1=0.9808342728297633,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.7077148337848485e-05

```

```

step: 10, loss: 9.284060070058331e-05
step: 20, loss: 7.282695878529921e-05
step: 30, loss: 4.623569475370459e-05
step: 40, loss: 0.08867929875850677
step: 50, loss: 6.0266873333603144e-05
step: 60, loss: 0.0002721527707763016
step: 70, loss: 3.7397014239104465e-05
step: 80, loss: 5.056061127106659e-05
step: 90, loss: 4.997327778255567e-05
step: 100, loss: 0.00030364730628207326
step: 110, loss: 0.0006172626162879169
step: 120, loss: 2.364021747780498e-05
step: 130, loss: 4.1240557038690895e-05
step: 140, loss: 3.6573670513462275e-05
step: 150, loss: 4.646456363843754e-05
step: 160, loss: 0.0004200846015010029
step: 170, loss: 3.293419285910204e-05
step: 180, loss: 4.75426932098344e-05
step: 190, loss: 5.306049934006296e-05
step: 200, loss: 3.127658783341758e-05
step: 210, loss: 4.941636507282965e-05
step: 220, loss: 3.4096250601578504e-05
step: 230, loss: 0.006280047353357077
/usr/local/lib/python3.7/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=0.9763779527559054, f1=0.9808773903262092,
best_f1=0.9761634506242906
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.947753637272399e-05
step: 10, loss: 2.746592326730024e-05
step: 20, loss: 3.2766754884505644e-05
step: 30, loss: 0.00011404320684960112
step: 40, loss: 6.555276195285842e-05
step: 50, loss: 0.011572490446269512
step: 60, loss: 0.0001118622676585801
step: 70, loss: 0.0002069891634164378
step: 80, loss: 0.0014323790092021227
step: 90, loss: 6.652955198660493e-05
step: 100, loss: 5.696090738638304e-05
step: 110, loss: 2.7763917387346737e-05
step: 120, loss: 0.000411814748076722
step: 130, loss: 3.834271410596557e-05
step: 140, loss: 5.076561865280382e-05
step: 150, loss: 7.183818524936214e-05
step: 160, loss: 2.7070916985394433e-05

```

```

step: 170, loss: 4.14801761507988e-05
step: 180, loss: 0.0009572265553288162
step: 190, loss: 0.00023142009740695357
step: 200, loss: 7.183517300290987e-05
step: 210, loss: 0.0007025222876109183
step: 220, loss: 6.986431253608316e-05
step: 230, loss: 9.291228343499824e-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=0.9753363228699552, f1=0.9809203142536477,
best_f1=0.9761634506242906
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, 220.37it/s]
load_f1 = 0.9764837625979844
real_f1 = 0.9730941704035874
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 [00:15, 274.91it/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 bert-base-uncased \
  --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 bert-base-uncased \
  --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: 0.612539529800415
step: 10, loss: 0.6007376909255981
step: 20, loss: 0.5624711513519287
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.2437007576227188
step: 40, loss: 0.32185128331184387
step: 50, loss: 0.18242985010147095
step: 60, loss: 0.05959119647741318
step: 70, loss: 0.13986727595329285
step: 80, loss: 0.081941619515419
step: 90, loss: 0.29854947328567505
step: 100, loss: 0.08296584337949753
step: 110, loss: 0.08092795312404633
step: 120, loss: 0.13931117951869965
step: 130, loss: 0.1499040573835373
step: 140, loss: 0.09027031809091568
step: 150, loss: 0.03967338055372238
step: 160, loss: 0.018117673695087433
step: 170, loss: 0.22706809639930725
step: 180, loss: 0.2165350317955017

```

```

step: 190, loss: 0.055701710283756256
step: 200, loss: 0.17940029501914978
step: 210, loss: 0.0318361334502697
step: 220, loss: 0.21765488386154175
step: 230, loss: 0.1720486879348755
step: 240, loss: 0.049955546855926514
step: 250, loss: 0.016379330307245255
step: 260, loss: 0.21765612065792084
step: 270, loss: 0.012840325012803078
step: 280, loss: 0.0360676534473896
step: 290, loss: 0.12615269422531128
step: 300, loss: 0.031007282435894012
step: 310, loss: 0.07493545860052109
step: 320, loss: 0.12486805766820908
step: 330, loss: 0.0275421142578125
step: 340, loss: 0.044410545378923416
step: 350, loss: 0.014216299168765545
step: 360, loss: 0.16979876160621643
step: 370, loss: 0.09490679204463959
step: 380, loss: 0.048034049570560455
step: 390, loss: 0.2706860303878784
step: 400, loss: 0.2467442750930786
step: 410, loss: 0.06360463798046112
step: 420, loss: 0.044862113893032074
step: 430, loss: 0.24631854891777039
step: 440, loss: 0.029950561001896858
step: 450, loss: 0.035796087235212326
step: 460, loss: 0.01994735561311245
step: 470, loss: 0.18450325727462769
step: 480, loss: 0.0899742990732193
step: 490, loss: 0.09610113501548767
step: 500, loss: 0.03191032260656357
step: 510, loss: 0.11166496574878693
step: 520, loss: 0.17629078030586243
step: 530, loss: 0.008954904973506927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9279404927940492, f1=0.9280373831775701,
best_f1=0.9280373831775701
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1695973128080368

```

step: 10, loss: 0.06522028893232346
step: 20, loss: 0.03267551586031914
step: 30, loss: 0.03492056205868721
step: 40, loss: 0.05062995105981827
step: 50, loss: 0.12876221537590027
step: 60, loss: 0.009125510230660439
step: 70, loss: 0.1287851780653
step: 80, loss: 0.04045557603240013
step: 90, loss: 0.012581935152411461
step: 100, loss: 0.08457966893911362
step: 110, loss: 0.011520937085151672
step: 120, loss: 0.14905765652656555
step: 130, loss: 0.13462843000888824
step: 140, loss: 0.033239495009183884
step: 150, loss: 0.060535840690135956
step: 160, loss: 0.044420093297958374
step: 170, loss: 0.03516561537981033
step: 180, loss: 0.02739412523806095
step: 190, loss: 0.061443135142326355
step: 200, loss: 0.03034185618162155
step: 210, loss: 0.04483763128519058
step: 220, loss: 0.14928443729877472
step: 230, loss: 0.013972745276987553
step: 240, loss: 0.011970628052949905
step: 250, loss: 0.1051458865404129
step: 260, loss: 0.006479908712208271
step: 270, loss: 0.24318918585777283
step: 280, loss: 0.006205155048519373
step: 290, loss: 0.06427910923957825
step: 300, loss: 0.1797700673341751
step: 310, loss: 0.046685829758644104
step: 320, loss: 0.15137533843517303
step: 330, loss: 0.0299531240016222
step: 340, loss: 0.0649411529302597
step: 350, loss: 0.0037367248442023993
step: 360, loss: 0.08548375219106674
step: 370, loss: 0.10999441146850586
step: 380, loss: 0.03665752336382866
step: 390, loss: 0.04773888364434242
step: 400, loss: 0.05775545537471771
step: 410, loss: 0.02215137705206871
step: 420, loss: 0.266801655292511
step: 430, loss: 0.028185002505779266
step: 440, loss: 0.16229966282844543
step: 450, loss: 0.033191580325365067
step: 460, loss: 0.015391351655125618
step: 470, loss: 0.023455481976270676
step: 480, loss: 0.193501815199852

```

step: 490, loss: 0.01025659404695034
step: 500, loss: 0.31333959102630615
step: 510, loss: 0.018794545903801918
step: 520, loss: 0.10907261818647385
step: 530, loss: 0.011690504848957062
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9404157043879908, f1=0.9320388349514563,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07547207921743393
step: 10, loss: 0.10974660515785217
step: 20, loss: 0.043557241559028625
step: 30, loss: 0.14370006322860718
step: 40, loss: 0.014082224108278751
step: 50, loss: 0.1122189462184906
step: 60, loss: 0.027691636234521866
step: 70, loss: 0.00854448787868023
step: 80, loss: 0.01642853207886219
step: 90, loss: 0.0032316267024725676
step: 100, loss: 0.038648881018161774
step: 110, loss: 0.013810219243168831
step: 120, loss: 0.09204892814159393
step: 130, loss: 0.008380815386772156
step: 140, loss: 0.08418113738298416
step: 150, loss: 0.12392687797546387
step: 160, loss: 0.04033222049474716
step: 170, loss: 0.17452390491962433
step: 180, loss: 0.06015633046627045
step: 190, loss: 0.10004155337810516
step: 200, loss: 0.06488199532032013
step: 210, loss: 0.04420146718621254
step: 220, loss: 0.020940491929650307
step: 230, loss: 0.054696280509233475
step: 240, loss: 0.0018130383687093854
step: 250, loss: 0.051420945674180984
step: 260, loss: 0.026890723034739494
step: 270, loss: 0.003982358146458864
step: 280, loss: 0.05497172474861145
step: 290, loss: 0.003442977322265506
step: 300, loss: 0.08986035734415054

```



```

step: 310, loss: 0.006005724426358938
step: 320, loss: 0.01726338267326355
step: 330, loss: 0.0027898463886231184
step: 340, loss: 0.01641993597149849
step: 350, loss: 0.0451357439160347
step: 360, loss: 0.018638119101524353
step: 370, loss: 0.007747657597064972
step: 380, loss: 0.01147131621837616
step: 390, loss: 0.02235812321305275
step: 400, loss: 0.008752563036978245
step: 410, loss: 0.004960167687386274
step: 420, loss: 0.10696311295032501
step: 430, loss: 0.014440351165831089
step: 440, loss: 0.009818029589951038
step: 450, loss: 0.035812657326459885
step: 460, loss: 0.01569502055644989
step: 470, loss: 0.03596826270222664
step: 480, loss: 0.020908020436763763
step: 490, loss: 0.00670167850330472
step: 500, loss: 0.037339624017477036
step: 510, loss: 0.004730351269245148
step: 520, loss: 0.131903275847435
step: 530, loss: 0.19350086152553558
/usr/local/lib/python3.7/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=0.9346136259716507, f1=0.9294492489758761,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023882290348410606
step: 10, loss: 0.0036671168636530638
step: 20, loss: 0.0022179584484547377
step: 30, loss: 0.01243915967643261
step: 40, loss: 0.002182109048590064
step: 50, loss: 0.01975908875465393
step: 60, loss: 0.02441808208823204
step: 70, loss: 0.0024599116295576096
step: 80, loss: 0.072407066822052
step: 90, loss: 0.11537441611289978
step: 100, loss: 0.011206711642444134
step: 110, loss: 0.010767092928290367
step: 120, loss: 0.001123512047342956
step: 130, loss: 0.001602945732884109
step: 140, loss: 0.006343752145767212
step: 150, loss: 0.02727525494992733
step: 160, loss: 0.0707787573337555

```

```

step: 170, loss: 0.007252747192978859
step: 180, loss: 0.058642324060201645
step: 190, loss: 0.13143014907836914
step: 200, loss: 0.02655891701579094
step: 210, loss: 0.010799426585435867
step: 220, loss: 0.006556638050824404
step: 230, loss: 0.07151476293802261
step: 240, loss: 0.011864340864121914
step: 250, loss: 0.0042561376467347145
step: 260, loss: 0.002471138024702668
step: 270, loss: 0.006026002578437328
step: 280, loss: 0.10887911170721054
step: 290, loss: 0.030600707978010178
step: 300, loss: 0.0005110171041451395
step: 310, loss: 0.001590306987054646
step: 320, loss: 0.03655937314033508
step: 330, loss: 0.00473359739407897
step: 340, loss: 0.0029972896445542574
step: 350, loss: 0.0748758465051651
step: 360, loss: 0.05884779989719391
step: 370, loss: 0.020837081596255302
step: 380, loss: 0.005481378640979528
step: 390, loss: 0.005005761049687862
step: 400, loss: 0.0018449736526235938
step: 410, loss: 0.007748785428702831
step: 420, loss: 0.03731592372059822
step: 430, loss: 0.04198446124792099
step: 440, loss: 0.00837025884538889
step: 450, loss: 0.019667528569698334
step: 460, loss: 0.039593856781721115
step: 470, loss: 0.11703465133905411
step: 480, loss: 0.027191834524273872
step: 490, loss: 0.004428436979651451
step: 500, loss: 0.03667779639363289
step: 510, loss: 0.02123061567544937
step: 520, loss: 0.03823247179389
step: 530, loss: 0.04972797632217407
/usr/local/lib/python3.7/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=0.9330166270783848, f1=0.9199999999999999,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.058525823056697845
step: 10, loss: 0.03690040856599808
step: 20, loss: 0.006459630560129881

```

step: 30, loss: 0.001638609217479825
step: 40, loss: 0.023685086518526077
step: 50, loss: 0.0008471276378259063
step: 60, loss: 0.0016830350505188107
step: 70, loss: 0.006774791516363621
step: 80, loss: 0.017636718228459358
step: 90, loss: 0.020267294719815254
step: 100, loss: 0.048251453787088394
step: 110, loss: 0.011666390113532543
step: 120, loss: 0.011005318723618984
step: 130, loss: 0.0016487633110955358
step: 140, loss: 0.01432145107537508
step: 150, loss: 0.03139529377222061
step: 160, loss: 0.0006563981878571212
step: 170, loss: 0.042963068932294846
step: 180, loss: 0.008547229692339897
step: 190, loss: 0.022387715056538582
step: 200, loss: 0.005698881112039089
step: 210, loss: 0.0517289936542511
step: 220, loss: 0.008221532218158245
step: 230, loss: 0.008512409403920174
step: 240, loss: 0.016249017789959908
step: 250, loss: 0.0007711664657108486
step: 260, loss: 0.003290983848273754
step: 270, loss: 0.0023546444717794657
step: 280, loss: 0.010318230837583542
step: 290, loss: 0.19981050491333008
step: 300, loss: 0.004178183153271675
step: 310, loss: 0.04586195945739746
step: 320, loss: 0.08877512067556381
step: 330, loss: 0.18828926980495453
step: 340, loss: 0.0021946895867586136
step: 350, loss: 0.0017443675315007567
step: 360, loss: 0.021190078929066658
step: 370, loss: 0.024897852912545204
step: 380, loss: 0.0009271949529647827
step: 390, loss: 0.0003172223223373294
step: 400, loss: 0.0359068401157856
step: 410, loss: 0.006764169316738844
step: 420, loss: 0.032952312380075455
step: 430, loss: 0.1309361308813095
step: 440, loss: 0.006775474641472101
step: 450, loss: 0.0034837096463888884
step: 460, loss: 0.003830639412626624
step: 470, loss: 0.006725902669131756
step: 480, loss: 0.006309840362519026
step: 490, loss: 0.0011841715313494205
step: 500, loss: 0.007080774288624525

```

step: 510, loss: 0.05169985443353653
step: 520, loss: 0.0006964142667129636
step: 530, loss: 0.002298937411978841
/usr/local/lib/python3.7/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=0.9233644859813084, f1=0.9222170470423847,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.043550390750169754
step: 10, loss: 0.0011228818912059069
step: 20, loss: 0.11631005257368088
step: 30, loss: 0.0028982805088162422
step: 40, loss: 0.012512968853116035
step: 50, loss: 0.008690607734024525
step: 60, loss: 0.0003205888206139207
step: 70, loss: 0.0032731303945183754
step: 80, loss: 0.0026876763440668583
step: 90, loss: 0.001203829189762473
step: 100, loss: 0.03198922798037529
step: 110, loss: 0.0027945549227297306
step: 120, loss: 0.0018240209901705384
step: 130, loss: 0.000325516244629398
step: 140, loss: 0.026325562968850136
step: 150, loss: 0.01186356507241726
step: 160, loss: 0.00112153182271868
step: 170, loss: 0.0009730765596032143
step: 180, loss: 0.0009958522859960794
step: 190, loss: 0.0008913589990697801
step: 200, loss: 0.0018787285080179572
step: 210, loss: 0.004296950995922089
step: 220, loss: 0.005003493744879961
step: 230, loss: 0.001644596690312028
step: 240, loss: 0.004812891129404306
step: 250, loss: 0.0005202738102525473
step: 260, loss: 0.0013772986130788922
step: 270, loss: 0.06291315704584122
step: 280, loss: 0.007486633490771055
step: 290, loss: 0.0001378756860503927
step: 300, loss: 0.0022486792877316475
step: 310, loss: 0.004352970514446497
step: 320, loss: 0.003070610575377941
step: 330, loss: 0.002423195168375969
step: 340, loss: 0.11394134908914566
step: 350, loss: 0.0005456984508782625
step: 360, loss: 0.0228460393846035

```

```
step: 370, loss: 0.09782504290342331
step: 380, loss: 0.0003547055821400136
step: 390, loss: 0.02618803083896637
step: 400, loss: 0.1689712405204773
step: 410, loss: 0.0012861844152212143
step: 420, loss: 0.0096141891553998
step: 430, loss: 0.003378876717761159
step: 440, loss: 0.0005865413113497198
step: 450, loss: 0.007902524434030056
step: 460, loss: 0.00018095631094183773
step: 470, loss: 0.015973227098584175
step: 480, loss: 0.0005924727884121239
step: 490, loss: 0.011510507203638554
step: 500, loss: 0.0004400070174597204
step: 510, loss: 0.0055885459296405315
step: 520, loss: 0.027672715485095978
step: 530, loss: 0.009249089285731316
/usr/local/lib/python3.7/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=0.9314179796107508, f1=0.9303184125519152,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006428018677979708
step: 10, loss: 0.0017036960925906897
step: 20, loss: 0.00035262989695183933
step: 30, loss: 0.045962147414684296
step: 40, loss: 0.16558726131916046
step: 50, loss: 0.01675632782280445
step: 60, loss: 0.007743889931589365
step: 70, loss: 0.0005883038393221796
step: 80, loss: 0.0017266873037442565
step: 90, loss: 0.0005696297739632428
step: 100, loss: 0.0007586744613945484
step: 110, loss: 0.0016795392148196697
step: 120, loss: 0.0021792433690279722
step: 130, loss: 0.005687952972948551
step: 140, loss: 0.002030841074883938
step: 150, loss: 0.006925404537469149
step: 160, loss: 0.0017430073348805308
step: 170, loss: 0.016240235418081284
step: 180, loss: 0.002093693008646369
step: 190, loss: 0.0016912742285057902
step: 200, loss: 0.0004905050154775381
step: 210, loss: 0.0023753882851451635
step: 220, loss: 0.0008328315452672541
```

```

step: 230, loss: 0.03464220091700554
step: 240, loss: 0.0005091334460303187
step: 250, loss: 0.00030671738204546273
step: 260, loss: 0.00013957842020317912
step: 270, loss: 0.002002979861572385
step: 280, loss: 0.001779036014340818
step: 290, loss: 0.0006181359058246017
step: 300, loss: 0.00474225590005517
step: 310, loss: 0.00045756157487630844
step: 320, loss: 0.0003393776423763484
step: 330, loss: 0.0016053059371188283
step: 340, loss: 0.033139102160930634
step: 350, loss: 0.00046437772107310593
step: 360, loss: 0.0012241447111591697
step: 370, loss: 0.0016049838159233332
step: 380, loss: 0.0014304801588878036
step: 390, loss: 0.0011242193868383765
step: 400, loss: 0.0012514443369582295
step: 410, loss: 0.00887310691177845
step: 420, loss: 0.00045626694918610156
step: 430, loss: 8.366588008357212e-05
step: 440, loss: 0.0007609363528899848
step: 450, loss: 0.003916352055966854
step: 460, loss: 0.0007729401113465428
step: 470, loss: 0.0056147328577935696
step: 480, loss: 0.048433106392621994
step: 490, loss: 0.02228441834449768
step: 500, loss: 0.002294318052008748
step: 510, loss: 0.020011894404888153
step: 520, loss: 0.013712609186768532
step: 530, loss: 0.03646249696612358
/usr/local/lib/python3.7/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=0.9309534992954438, f1=0.9283054003724395,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006287486292421818
step: 10, loss: 0.022284941747784615
step: 20, loss: 0.00017863484390545636
step: 30, loss: 0.01746944896876812
step: 40, loss: 0.0003265242849010974
step: 50, loss: 0.009374422952532768
step: 60, loss: 0.0044149667955935
step: 70, loss: 0.0002858213265426457
step: 80, loss: 0.004177585709840059

```

step: 90, loss: 0.0015793804777786136
step: 100, loss: 0.0006096781580708921
step: 110, loss: 0.0014509472530335188
step: 120, loss: 0.001545475795865059
step: 130, loss: 0.00028771947836503386
step: 140, loss: 0.0018895177636295557
step: 150, loss: 8.738039468880743e-05
step: 160, loss: 0.002819651272147894
step: 170, loss: 0.1573459953069687
step: 180, loss: 0.0015921695157885551
step: 190, loss: 0.0014021472306922078
step: 200, loss: 0.0024819679092615843
step: 210, loss: 0.010667751543223858
step: 220, loss: 0.02137145958840847
step: 230, loss: 0.00024570507230237126
step: 240, loss: 0.002168073086068034
step: 250, loss: 0.007616675458848476
step: 260, loss: 0.00035161798587068915
step: 270, loss: 0.06416121125221252
step: 280, loss: 0.00951263215392828
step: 290, loss: 0.0004756729758810252
step: 300, loss: 0.0006635561003349721
step: 310, loss: 0.00012914426042698324
step: 320, loss: 0.002603285713121295
step: 330, loss: 0.00208613695576787
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 340, loss: 7.458896288881078e-05
step: 350, loss: 0.03453119099140167
step: 360, loss: 0.001095122192054987
step: 370, loss: 0.0001761282910592854
step: 380, loss: 0.00027704835520125926
step: 390, loss: 0.1439063847064972
step: 400, loss: 0.0007303737802430987
step: 410, loss: 0.001799126504920423
step: 420, loss: 0.0018453693483024836
step: 430, loss: 0.016751756891608238
step: 440, loss: 0.00018392868514638394
step: 450, loss: 0.0001745184708852321
step: 460, loss: 0.0014098670799285173
step: 470, loss: 0.0009314129711128771
step: 480, loss: 0.006192260887473822
step: 490, loss: 0.0013981384690850973
step: 500, loss: 0.003034073393791914
step: 510, loss: 0.0034522339701652527
step: 520, loss: 0.0035228286869823933
step: 530, loss: 0.0002499047259334475
/usr/local/lib/python3.7/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=0.9337042188224385, f1=0.9278445883441258,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017573704244568944
step: 10, loss: 0.009481282904744148
step: 20, loss: 0.013248326256871223
step: 30, loss: 0.00044694339158013463
step: 40, loss: 0.00043145116069354117
step: 50, loss: 0.0005805812543258071
step: 60, loss: 0.0001743226166581735
step: 70, loss: 0.0066195297986269
step: 80, loss: 0.003220738610252738
step: 90, loss: 0.0004192916676402092
step: 100, loss: 0.012838626280426979
step: 110, loss: 0.0007751917582936585
step: 120, loss: 0.010284420102834702
step: 130, loss: 0.006471442058682442
step: 140, loss: 0.1012495905160904
step: 150, loss: 0.00025465694488957524
step: 160, loss: 0.001144660054706037
step: 170, loss: 0.012486564926803112
step: 180, loss: 0.003937527537345886
step: 190, loss: 0.04122256487607956
step: 200, loss: 0.000624787644483149
step: 210, loss: 0.0012394190998747945
step: 220, loss: 0.005510449409484863
step: 230, loss: 0.0005243556224741042
step: 240, loss: 0.0013435790315270424
step: 250, loss: 0.0005232792464084923
step: 260, loss: 0.0034960173070430756
step: 270, loss: 0.0021423869766294956
step: 280, loss: 0.0027560784947127104
step: 290, loss: 0.021110357716679573
step: 300, loss: 8.954948862083256e-05
step: 310, loss: 0.005259099882096052
step: 320, loss: 9.64014689088799e-05
step: 330, loss: 0.00028351915534585714
step: 340, loss: 0.01096726581454277
step: 350, loss: 0.10027226060628891
step: 360, loss: 0.00029067715513519943
step: 370, loss: 0.0007017283933237195
step: 380, loss: 0.0005643905606120825
step: 390, loss: 0.0015063855098560452
step: 400, loss: 0.00250408542342484
step: 410, loss: 0.00021554350678343326
```



```
step: 420, loss: 0.00021258114429656416
step: 430, loss: 0.00022444178466685116
step: 440, loss: 0.0010923697846010327
step: 450, loss: 0.00014773134898860008
step: 460, loss: 0.00028959388146176934
step: 470, loss: 4.434067159309052e-05
step: 480, loss: 0.00010122691310243681
step: 490, loss: 0.025851821526885033
step: 500, loss: 0.002188492799177766
step: 510, loss: 0.07345735281705856
step: 520, loss: 0.0007180970278568566
step: 530, loss: 0.0025032127741724253
/usr/local/lib/python3.7/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=0.9324515824279641, f1=0.9255419415645617,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0009919940494000912
step: 10, loss: 0.00014846927660983056
step: 20, loss: 9.890851652016863e-05
step: 30, loss: 0.0041439286433160305
step: 40, loss: 9.627994586480781e-05
step: 50, loss: 0.00012456455442588776
step: 60, loss: 0.0004332124663051218
step: 70, loss: 0.0003782058774959296
step: 80, loss: 0.00011916742369066924
step: 90, loss: 6.115844735177234e-05
step: 100, loss: 0.0009292068425565958
step: 110, loss: 0.00024488335475325584
step: 120, loss: 7.39885144867003e-05
step: 130, loss: 0.00030081780278123915
step: 140, loss: 0.0035655044484883547
step: 150, loss: 0.026545992121100426
step: 160, loss: 0.04385937750339508
step: 170, loss: 0.0001267561747226864
step: 180, loss: 0.0005483294371515512
step: 190, loss: 0.0016057504108175635
step: 200, loss: 0.0022350852377712727
step: 210, loss: 0.004507541656494141
step: 220, loss: 0.013093451969325542
step: 230, loss: 0.0007459177286364138
step: 240, loss: 0.001497928868047893
step: 250, loss: 7.888372056186199e-05
step: 260, loss: 0.0018091417150571942
step: 270, loss: 0.011634387075901031
```

```
step: 280, loss: 0.0002541830763220787
step: 290, loss: 0.0015062065795063972
step: 300, loss: 3.515453863656148e-05
step: 310, loss: 0.0025460466276854277
step: 320, loss: 0.0075770178809762
step: 330, loss: 0.000369425630196929
step: 340, loss: 0.0017335321754217148
step: 350, loss: 0.0003181329811923206
step: 360, loss: 0.00022517589968629181
step: 370, loss: 0.003213524119928479
step: 380, loss: 0.005196982063353062
step: 390, loss: 0.0007015284500084817
step: 400, loss: 0.0008512192871421576
step: 410, loss: 0.0015464364551007748
step: 420, loss: 0.0006984236533753574
step: 430, loss: 0.0001343133335467428
step: 440, loss: 0.0002944017469417304
step: 450, loss: 6.513280823128298e-05
step: 460, loss: 0.0004636655794456601
step: 470, loss: 0.00020512212358880788
step: 480, loss: 0.009149348363280296
step: 490, loss: 9.555427823215723e-05
step: 500, loss: 0.011778326705098152
step: 510, loss: 0.005636685527861118
step: 520, loss: 0.0004733378009404987
step: 530, loss: 0.0003338250098749995
/usr/local/lib/python3.7/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=0.9284043051006083, f1=0.9245810055865922,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006342533743008971
step: 10, loss: 0.0078055644407868385
step: 20, loss: 0.0005781407235190272
step: 30, loss: 0.00038076608325354755
step: 40, loss: 0.02490786463022232
step: 50, loss: 0.0009741586982272565
step: 60, loss: 0.020133057609200478
step: 70, loss: 0.001086935168132186
step: 80, loss: 0.000649516296107322
step: 90, loss: 9.484904148848727e-05
step: 100, loss: 0.006409071385860443
step: 110, loss: 0.0025458454620093107
step: 120, loss: 5.5212705774465576e-05
step: 130, loss: 5.3545369155472144e-05
```

```

step: 140, loss: 0.00043154251761734486
step: 150, loss: 0.0006277438369579613
step: 160, loss: 0.00037447086651809514
step: 170, loss: 0.00042942914296872914
step: 180, loss: 0.0003211272123735398
step: 190, loss: 0.002582363085821271
step: 200, loss: 0.00443987175822258
step: 210, loss: 0.007273130584508181
step: 220, loss: 0.0007280074059963226
step: 230, loss: 0.00018236774485558271
step: 240, loss: 0.000370649213436991
step: 250, loss: 3.476162601145916e-05
step: 260, loss: 5.055178189650178e-05
step: 270, loss: 0.0015917748678475618
step: 280, loss: 0.002789123449474573
step: 290, loss: 0.0010150392772629857
step: 300, loss: 4.595027712639421e-05
step: 310, loss: 0.00023889644944574684
step: 320, loss: 2.380047044425737e-05
step: 330, loss: 4.179299503448419e-05
step: 340, loss: 7.231304334709421e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 350, loss: 0.0017985226586461067
step: 360, loss: 0.00021044487948529422
step: 370, loss: 0.00030319119105115533
step: 380, loss: 0.00018445003661327064
step: 390, loss: 2.4787568690953776e-05
step: 400, loss: 0.00012688741844613105
step: 410, loss: 0.002755097346380353
step: 420, loss: 7.081671355990693e-05
step: 430, loss: 1.7907246729009785e-05
step: 440, loss: 0.0006318610976450145
step: 450, loss: 6.906757334945723e-05
step: 460, loss: 0.008290276862680912
step: 470, loss: 0.0002813223691191524
step: 480, loss: 0.0007952692685648799
step: 490, loss: 0.0007242323481477797
step: 500, loss: 0.00013028782268520445
step: 510, loss: 3.222544546588324e-05
step: 520, loss: 2.227643744845409e-05
step: 530, loss: 0.004470885265618563
/usr/local/lib/python3.7/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=0.9322671683913453, f1=0.9284386617100371,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.208261766005307e-05
step: 10, loss: 2.2932143110665493e-05
step: 20, loss: 6.314588245004416e-05
step: 30, loss: 0.005377018358558416
step: 40, loss: 0.0002304660010850057
step: 50, loss: 0.10306071490049362
step: 60, loss: 0.0024136349093168974
step: 70, loss: 0.0002352394221816212
step: 80, loss: 0.00012763863196596503
step: 90, loss: 2.688017593754921e-05
step: 100, loss: 0.00014588648627977818
step: 110, loss: 0.0007943774689920247
step: 120, loss: 0.00010119465150637552
step: 130, loss: 0.00035081591340713203
step: 140, loss: 2.6799074476002716e-05
step: 150, loss: 3.7293302739271894e-05
step: 160, loss: 3.2387557439506054e-05
step: 170, loss: 0.0006306211580522358
step: 180, loss: 7.54933716962114e-05
step: 190, loss: 0.0006652857409790158
step: 200, loss: 0.00023623008746653795
step: 210, loss: 0.00039329382707364857
step: 220, loss: 7.785351772326976e-05
step: 230, loss: 3.982018097303808e-05
step: 240, loss: 0.0006356038502417505
step: 250, loss: 9.886796033242717e-05
step: 260, loss: 0.00012875613174401224
step: 270, loss: 4.9462727474747226e-05
step: 280, loss: 0.00032613053917884827
step: 290, loss: 0.00016292027430608869
step: 300, loss: 0.15391884744167328
step: 310, loss: 0.0005761931533925235
step: 320, loss: 0.0002674439747352153
step: 330, loss: 0.00044514398905448616
step: 340, loss: 0.003122118767350912
step: 350, loss: 0.0002823449904099107
step: 360, loss: 0.0007229824550449848
step: 370, loss: 0.003027458442375064
step: 380, loss: 0.00030347079155035317
step: 390, loss: 0.0016159974038600922
step: 400, loss: 0.0006729810265824199
step: 410, loss: 0.0037917979061603546
step: 420, loss: 0.006473610643297434
step: 430, loss: 0.0037030442617833614
step: 440, loss: 0.0007074567256495357
step: 450, loss: 0.004100136924535036
step: 460, loss: 0.00015446910401806235
```

```

step: 470, loss: 0.00014502473641186953
step: 480, loss: 0.009742665104568005
step: 490, loss: 0.041484285145998
step: 500, loss: 2.8106111130909994e-05
step: 510, loss: 0.17260821163654327
step: 520, loss: 5.0213020585943013e-05
step: 530, loss: 2.0745384972542524e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9325210871602625, f1=0.9307298930729893,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032883891835808754
step: 10, loss: 1.7974294678424485e-05
step: 20, loss: 5.9062029322376475e-05
step: 30, loss: 0.003215006086975336
step: 40, loss: 4.874963269685395e-05
step: 50, loss: 0.00014905465650372207
step: 60, loss: 6.011091682012193e-05
step: 70, loss: 5.857169162482023e-05
step: 80, loss: 3.5291872336529195e-05
step: 90, loss: 0.008073651231825352
step: 100, loss: 0.0008895514765754342
step: 110, loss: 0.0004022720968350768
step: 120, loss: 6.966797809582204e-05
step: 130, loss: 4.066131077706814e-05
step: 140, loss: 0.0004455853486433625
step: 150, loss: 0.0002519389963708818
step: 160, loss: 2.717023562581744e-05
step: 170, loss: 0.00010492293222341686
step: 180, loss: 6.648853741353378e-05
step: 190, loss: 0.00013978377683088183
step: 200, loss: 8.338138286489993e-05
step: 210, loss: 0.00042107916669920087
step: 220, loss: 2.1788806407130323e-05
step: 230, loss: 0.0009437822154723108
step: 240, loss: 0.0014084193389862776
step: 250, loss: 4.705788160208613e-05
step: 260, loss: 0.009297044947743416
step: 270, loss: 3.383804869372398e-05
step: 280, loss: 0.0005144180031493306
step: 290, loss: 0.0002941979328170419
step: 300, loss: 8.141156286001205e-05
step: 310, loss: 0.0010307912016287446
step: 320, loss: 0.0019692375790327787

```

```

step: 330, loss: 0.029122615233063698
step: 340, loss: 0.0010816220892593265
step: 350, loss: 6.48352870484814e-05
step: 360, loss: 0.0006878736312501132
step: 370, loss: 0.00010189467866439372
step: 380, loss: 3.7116973544470966e-05
step: 390, loss: 0.0004993322654627264
step: 400, loss: 0.00013613600458484143
step: 410, loss: 0.0009318387019447982
step: 420, loss: 3.90896930184681e-05
step: 430, loss: 0.0001875103043857962
step: 440, loss: 0.0043356819078326225
step: 450, loss: 0.0006321895634755492
step: 460, loss: 0.001306155463680625
step: 470, loss: 0.0003849720233120024
step: 480, loss: 8.811564475763589e-05
step: 490, loss: 0.0003565190127119422
step: 500, loss: 0.00017976530943997204
step: 510, loss: 0.0005412086029537022
step: 520, loss: 0.0023307872470468283
step: 530, loss: 0.00010679599654395133
/usr/local/lib/python3.7/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=0.9321642824180896, f1=0.9321642824180896,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019874349236488342
step: 10, loss: 3.432609446463175e-05
step: 20, loss: 0.00026155277737416327
step: 30, loss: 0.00011183810420334339
step: 40, loss: 0.006163571495562792
step: 50, loss: 0.003831228706985712
step: 60, loss: 1.2226294529682491e-05
step: 70, loss: 0.0005673042614944279
step: 80, loss: 2.1788251615362242e-05
step: 90, loss: 0.00010023025242844597
step: 100, loss: 0.00044856162276118994
step: 110, loss: 0.0001318799622822553
step: 120, loss: 0.0003291511966381222
step: 130, loss: 0.010910570621490479
step: 140, loss: 0.0002275221486343071
step: 150, loss: 0.0002035423822235316
step: 160, loss: 0.0002912736963480711
step: 170, loss: 4.522491872194223e-05
step: 180, loss: 3.3131593227153644e-05

```

```

step: 190, loss: 4.165831342106685e-05
step: 200, loss: 0.002237205859273672
step: 210, loss: 2.6142508431803435e-05
step: 220, loss: 0.0005520968115888536
step: 230, loss: 0.003382098861038685
step: 240, loss: 4.035193342133425e-05
step: 250, loss: 0.00010394854325568303
step: 260, loss: 0.0034223017282783985
step: 270, loss: 0.00032291028765030205
step: 280, loss: 2.9241868105600588e-05
step: 290, loss: 0.0003777882957365364
step: 300, loss: 0.004657924175262451
step: 310, loss: 0.0003413849335629493
step: 320, loss: 0.0005653548869304359
step: 330, loss: 0.03386924788355827
step: 340, loss: 9.02196261449717e-05
step: 350, loss: 0.001506954780779779
step: 360, loss: 0.005661098752170801
step: 370, loss: 0.00013036544260103256
step: 380, loss: 0.0003216266050003469
step: 390, loss: 0.019820561632514
step: 400, loss: 0.0025441194884479046
step: 410, loss: 0.00028046660008840263
step: 420, loss: 0.017232252284884453
step: 430, loss: 0.059314992278814316
step: 440, loss: 0.1599215865135193
step: 450, loss: 0.0006028801435604692
step: 460, loss: 0.00014642317546531558
step: 470, loss: 0.000897340418305248
step: 480, loss: 0.004113788716495037
step: 490, loss: 6.718761869706213e-05
step: 500, loss: 0.00040303950663655996
step: 510, loss: 0.001590378349646926
step: 520, loss: 0.0005415552295744419
step: 530, loss: 0.002628063317388296
/usr/local/lib/python3.7/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=0.9294287085933749, f1=0.9243937232524965,
best_f1=0.9320388349514563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.3403113775420934e-05
step: 10, loss: 0.006396114826202393
step: 20, loss: 0.0008520123083144426
step: 30, loss: 8.180148870451376e-05
step: 40, loss: 0.0032774663995951414

```

step: 50, loss: 0.00632402254268527
step: 60, loss: 4.067784539074637e-05
step: 70, loss: 0.0006301904213614762
step: 80, loss: 0.00042377557838335633
step: 90, loss: 0.00015042908489704132
step: 100, loss: 8.846016862662509e-05
step: 110, loss: 0.00021637165627907962
step: 120, loss: 0.0012302040122449398
step: 130, loss: 0.06937138736248016
step: 140, loss: 5.214561315369792e-05
step: 150, loss: 0.017290500923991203
step: 160, loss: 9.694366599433124e-05
step: 170, loss: 8.130733476718888e-05
step: 180, loss: 3.340329567436129e-05
step: 190, loss: 0.0005319345509633422
step: 200, loss: 0.0002581071457825601
step: 210, loss: 0.0002659937890712172
step: 220, loss: 4.821053153136745e-05
step: 230, loss: 0.0007281076395884156
step: 240, loss: 7.217701931949705e-05
step: 250, loss: 0.00010663629655027762
step: 260, loss: 0.00011303027713438496
step: 270, loss: 2.0272649635444395e-05
step: 280, loss: 0.0002021140680881217
step: 290, loss: 3.8019174098735675e-05
step: 300, loss: 4.072340016136877e-05
step: 310, loss: 0.00022253874340094626
step: 320, loss: 0.0001415690203430131
step: 330, loss: 0.00033415108919143677
step: 340, loss: 0.0001441667991457507
step: 350, loss: 0.0032087310682982206
step: 360, loss: 0.004294856917113066
step: 370, loss: 8.021417306736112e-05
step: 380, loss: 6.675915938103572e-05
step: 390, loss: 0.013451694510877132
step: 400, loss: 5.010953827877529e-05
step: 410, loss: 0.0020385445095598698
step: 420, loss: 0.0003100468311458826
step: 430, loss: 0.0011648408835753798
step: 440, loss: 0.0005582162993960083
step: 450, loss: 9.734220657264814e-05
step: 460, loss: 2.604227302072104e-05
step: 470, loss: 0.00019324007735121995
step: 480, loss: 2.4090219085337594e-05
step: 490, loss: 0.0236648116260767
step: 500, loss: 0.00028630258748307824
step: 510, loss: 6.424028833862394e-05
step: 520, loss: 2.914550350396894e-05


```

step: 530, loss: 0.00014154837117530406
/usr/local/lib/python3.7/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=0.9291784702549575, f1=0.9312119794103885,
best_f1=0.9320388349514563
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:21, 266.15it/s]
load_f1 = 0.93666204345816
real_f1 = 0.9342592592592593
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 266.48it/s]

```