

BDirty_50_3_5_bert

August 13, 2022

0.1 Install packages

```
[ ]: !git clone https://github.com/megagonlabs/ditto
      %cd ditto
      !pip install -r requirements.txt
      !git clone https://github.com/rit-git/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 | 20.94 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 1.7 MB/s
Collecting numpy==1.19.2
  Downloading numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
    |                               | 14.5 MB 43.8 MB/s
Collecting regex==2019.12.20
  Downloading regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
    |                               | 689 kB 54.5 MB/s
Collecting scipy==1.3.2
  Downloading scipy-1.3.2-cp37-cp37m-manylinux1_x86_64.whl (25.2 MB)
    |                               | 25.2 MB 1.4 MB/s
Collecting sentencepiece==0.1.85
  Downloading sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
    |                               | 1.0 MB 58.0 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 57.7 MB/s
ERROR: Could not find a version that satisfies the requirement
torch==1.9.0+cu111 (from versions: 1.0.0, 1.0.1, 1.0.1.post2, 1.1.0, 1.2.0,
1.3.0, 1.3.1, 1.4.0, 1.5.0, 1.5.1, 1.6.0, 1.7.0, 1.7.1, 1.8.0, 1.8.1, 1.9.0,
1.9.1, 1.10.0, 1.10.1, 1.10.2, 1.11.0, 1.12.0, 1.12.1)
ERROR: No matching distribution found for torch==1.9.0+cu111
Cloning into 'Snippext_public'...
remote: Enumerating objects: 413, done.
remote: Counting objects: 100% (22/22), done.
remote: Compressing objects: 100% (8/8), done.
remote: Total 413 (delta 15), reused 14 (delta 14), pack-reused 391
Receiving objects: 100% (413/413), 21.06 MiB | 19.83 MiB/s, done.
Resolving deltas: 100% (196/196), done.
/content/ditto/Snippext_public
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting gensim==3.8.1
  Using cached gensim-3.8.1-cp37-cp37m-manylinux1_x86_64.whl (24.2 MB)
Collecting numpy==1.19.2
  Using cached numpy-1.19.2-cp37-cp37m-manylinux2010_x86_64.whl (14.5 MB)
Collecting regex==2019.12.20
  Using cached regex-2019.12.20-cp37-cp37m-manylinux2010_x86_64.whl (689 kB)
Collecting spacy==2.2.3
  Downloading spacy-2.2.3-cp37-cp37m-manylinux1_x86_64.whl (10.4 MB)
  | 10.4 MB 5.2 MB/s
Collecting sentencepiece==0.1.85
  Using cached sentencepiece-0.1.85-cp37-cp37m-manylinux1_x86_64.whl (1.0 MB)
Requirement already satisfied: sklearn==0.0 in /usr/local/lib/python3.7/dist-
packages (from -r requirements.txt (line 6)) (0.0)
Collecting tensorboardX==2.0
  Downloading tensorboardX-2.0-py2.py3-none-any.whl (195 kB)
  | 195 kB 68.5 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 5.3 MB/s
Collecting transformers==3.1.0
  Downloading transformers-3.1.0-py3-none-any.whl (884 kB)
  | 884 kB 55.4 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 25.9 MB/s
Requirement already satisfied: six>=1.5.0 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (1.15.0)
Requirement already satisfied: scipy>=0.18.1 in /usr/local/lib/python3.7/dist-
packages (from gensim==3.8.1->-r requirements.txt (line 1)) (1.7.3)
Requirement already satisfied: smart-open>=1.8.1 in
/usr/local/lib/python3.7/dist-packages (from gensim==3.8.1->-r requirements.txt
(line 1)) (5.2.1)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (1.0.7)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (2.0.6)
Collecting catalogue<1.1.0,>=0.0.7
    Downloading catalogue-1.0.0-py2.py3-none-any.whl (7.7 kB)
Collecting plac<1.2.0,>=0.9.6
    Downloading plac-1.1.3-py2.py3-none-any.whl (20 kB)
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: wasabi<1.1.0,>=0.4.0 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (0.10.1)
Collecting srsly<1.1.0,>=0.1.0
    Downloading srsly-1.0.5-cp37-cp37m-manylinux2014_x86_64.whl (184 kB)
      |                               | 184 kB 60.9 MB/s
Requirement already satisfied: setuptools in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (57.4.0)
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 57.9 MB/s
Collecting thinc<7.4.0,>=7.3.0
    Downloading thinc-7.3.1-cp37-cp37m-manylinux1_x86_64.whl (2.2 MB)
      |                               | 2.2 MB 49.3 MB/s
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in
/usr/local/lib/python3.7/dist-packages (from spacy==2.2.3->-r requirements.txt
(line 4)) (3.0.6)
Requirement already satisfied: scikit-learn in /usr/local/lib/python3.7/dist-
packages (from sklearn==0.0->-r requirements.txt (line 6)) (1.0.2)
Requirement already satisfied: protobuf>=3.8.0 in /usr/local/lib/python3.7/dist-
packages (from tensorboardX==2.0->-r requirements.txt (line 8)) (3.17.3)
Collecting sacremoses

```

```

    Downloading sacremoses-0.0.53.tar.gz (880 kB)
      |                               | 880 kB 56.8 MB/s
Collecting tokenizers==0.8.1.rc2
  Downloading tokenizers-0.8.1rc2-cp37-cp37m-manylinux1_x86_64.whl (3.0 MB)
      |                               | 3.0 MB 45.4 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: packaging in /usr/local/lib/python3.7/dist-packages (from transformers==3.1.0->-r requirements.txt (line 11)) (21.3)
Requirement already satisfied: importlib-metadata>=0.20 in /usr/local/lib/python3.7/dist-packages (from catalogue<1.1.0,>=0.0.7->spacy==2.2.3->-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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (2.10)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (3.0.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy==2.2.3->-r requirements.txt (line 4)) (1.24.3)
Requirement already satisfied: 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: 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=1449920 sha256=7b4e54358594e7a4c62d1a89aef40050aa40f3263c38bd2dd31a6ccad10f094b
    Stored in directory: /root/.cache/pip/wheels/48/8b/7f/473521e0c731c6566d631b281f323842bbda9bd819eb9a3ead

```

```

Building wheel for sacremoses (setup.py) ... done
Created wheel for sacremoses: filename=sacremoses-0.0.53-py3-none-any.whl
size=895260
sha256=47260f3b02135aeacc5076f53c8c05630e456e5269eb32995837f24b3be44533
Stored in directory: /root/.cache/pip/wheels/87/39/dd/a83eeef36d0bf98e7a4d1933
a4ad2d660295a40613079bafc9
Successfully built nltk sacremoses
Installing collected packages: numpy, tqdm, srsly, regex, plac, blis,
tokenizers, thinc, sentencepiece, sacremoses, catalogue, transformers, torch,
tensorboardX, spacy, nltk, jsonlines, gensim
Attempting uninstall: numpy
  Found existing installation: numpy 1.21.6
  Uninstalling numpy-1.21.6:
    Successfully uninstalled numpy-1.21.6
Attempting uninstall: tqdm
  Found existing installation: tqdm 4.64.0
  Uninstalling tqdm-4.64.0:
    Successfully uninstalled tqdm-4.64.0
Attempting uninstall: srsly
  Found existing installation: srsly 2.4.4
  Uninstalling srsly-2.4.4:
    Successfully uninstalled srsly-2.4.4
Attempting uninstall: regex
  Found existing installation: regex 2022.6.2
  Uninstalling regex-2022.6.2:
    Successfully uninstalled regex-2022.6.2
Attempting uninstall: blis
  Found existing installation: blis 0.7.8
  Uninstalling blis-0.7.8:
    Successfully uninstalled blis-0.7.8
Attempting uninstall: thinc
  Found existing installation: thinc 8.1.0
  Uninstalling thinc-8.1.0:
    Successfully uninstalled thinc-8.1.0
Attempting uninstall: catalogue
  Found existing installation: catalogue 2.0.8
  Uninstalling catalogue-2.0.8:
    Successfully uninstalled catalogue-2.0.8
Attempting uninstall: torch
  Found existing installation: torch 1.12.0+cu113
  Uninstalling torch-1.12.0+cu113:
    Successfully uninstalled torch-1.12.0+cu113
Attempting uninstall: spacy
  Found existing installation: spacy 3.4.1
  Uninstalling spacy-3.4.1:
    Successfully uninstalled spacy-3.4.1
Attempting uninstall: nltk
  Found existing installation: nltk 3.7

```

```

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

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

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

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

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

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

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

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

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

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

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

```

```
[ ]: True
```

0.2 Install fp16 optimization

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

```
Cloning into 'apex'...
remote: Enumerating objects: 10085, done.
remote: Counting objects: 100% (201/201), done.
remote: Compressing objects: 100% (138/138), done.
remote: Total 10085 (delta 98), reused 131 (delta 59), pack-reused 9884
Receiving objects: 100% (10085/10085), 14.95 MiB | 21.90 MiB/s, done.
Resolving deltas: 100% (6902/6902), done.
/content/ditto/apex
Using pip 21.1.3 from /usr/local/lib/python3.7/dist-packages/pip (python 3.7)
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Non-user install because site-packages writable
Created temporary directory: /tmp/pip-ephem-wheel-cache-vvkbei0i
Created temporary directory: /tmp/pip-req-tracker-c2evqbtv
Initialized build tracking at /tmp/pip-req-tracker-c2evqbtv
Created build tracker: /tmp/pip-req-tracker-c2evqbtv
Entered build tracker: /tmp/pip-req-tracker-c2evqbtv
```

```

Created temporary directory: /tmp/pip-install-3mtbnc68
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-n2r7qhrn
  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-c2evqbtv'
  Running setup.py (path:/tmp/pip-req-build-n2r7qhrn/setup.py) egg_info for
package from file:///content/ditto/apex
  Created temporary directory: /tmp/pip-pip-egg-info-8a62i9vr
  Running command python setup.py egg_info

  torch.__version__ = 1.4.0

  running egg_info
  creating /tmp/pip-pip-egg-info-8a62i9vr/apex.egg-info
  writing /tmp/pip-pip-egg-info-8a62i9vr/apex.egg-info/PKG-INFO
  writing dependency_links to /tmp/pip-pip-egg-info-8a62i9vr/apex.egg-
info/dependency_links.txt
  writing top-level names to /tmp/pip-pip-egg-info-8a62i9vr/apex.egg-
info/top_level.txt
  writing manifest file '/tmp/pip-pip-egg-info-8a62i9vr/apex.egg-
info/SOURCES.txt'
  adding license file 'LICENSE'
  writing manifest file '/tmp/pip-pip-egg-info-8a62i9vr/apex.egg-
info/SOURCES.txt'
  Source in /tmp/pip-req-build-n2r7qhrn 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-c2evqbtv'
Created temporary directory: /tmp/pip-unpack-_5sok9gg
Building wheels for collected packages: apex
  Created temporary directory: /tmp/pip-wheel-8lq2_k3d
  Building wheel for apex (setup.py) ... Destination directory:
/tmp/pip-wheel-8lq2_k3d
  Running command /usr/bin/python3 -u -c 'import io, os, sys, setuptools,
tokenize; sys.argv[0] = ''''/tmp/pip-req-build-n2r7qhrn/setup.py'''';

```



```
__file__='''/tmp/pip-req-build-n2r7qhrn/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-8lq2_k3d
```

```
torch.__version__ = 1.4.0
```

```
running bdist_wheel
running build
running build_py
creating build
creating build/lib
creating build/lib/apex
copying apex/__init__.py -> build/lib/apex
copying apex/_autocast_utils.py -> build/lib/apex
creating build/lib/apex/normalization
copying apex/normalization/__init__.py -> build/lib/apex/normalization
copying apex/normalization/fused_layer_norm.py -> build/lib/apex/normalization
creating build/lib/apex/mlp
copying apex/mlp/__init__.py -> build/lib/apex/mlp
copying apex/mlp/mlp.py -> build/lib/apex/mlp
creating build/lib/apex/amp
copying apex/amp/opt.py -> build/lib/apex/amp
copying apex/amp/scaler.py -> build/lib/apex/amp
copying apex/amp/frontend.py -> build/lib/apex/amp
copying apex/amp/rnn_compat.py -> build/lib/apex/amp
copying apex/amp/__version__.py -> build/lib/apex/amp
copying apex/amp/utils.py -> build/lib/apex/amp
copying apex/amp/__init__.py -> build/lib/apex/amp
copying apex/amp/amp.py -> build/lib/apex/amp
copying apex/amp/compat.py -> build/lib/apex/amp
copying apex/amp/wrap.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/_process_optimizer.py -> build/lib/apex/amp
copying apex/amp/handle.py -> build/lib/apex/amp
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/fused_dense
copying apex/fused_dense/__init__.py -> build/lib/apex/fused_dense
copying apex/fused_dense/fused_dense.py -> build/lib/apex/fused_dense
creating build/lib/apex/parallel
```

```

    copying apex/parallel/__init__.py -> build/lib/apex/parallel
    copying apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/lib/apex/parallel
    copying apex/parallel/multiproc.py -> build/lib/apex/parallel
    copying apex/parallel/sync_batchnorm.py -> build/lib/apex/parallel
    copying apex/parallel/LARC.py -> build/lib/apex/parallel
    copying apex/parallel/distributed.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
    creating 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_adam.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_mixed_precision_lamb.py ->
build/lib/apex/optimizers
    copying apex/optimizers/fused_adagrad.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_novograd.py -> build/lib/apex/optimizers
    copying apex/optimizers/fused_lamb.py -> build/lib/apex/optimizers
    creating build/lib/apex/fp16_utils
    copying apex/fp16_utils/fp16_optimizer.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/__init__.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/fp16util.py -> build/lib/apex/fp16_utils
    copying apex/fp16_utils/loss_scaler.py -> build/lib/apex/fp16_utils
    creating build/lib/apex/transformer
    copying apex/transformer/microbatches.py -> build/lib/apex/transformer
    copying apex/transformer/parallel_state.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/enums.py -> build/lib/apex/transformer
    copying apex/transformer/log_util.py -> build/lib/apex/transformer
    creating build/lib/apex/RNN
    copying apex/RNN/cells.py -> build/lib/apex/RNN
    copying apex/RNN/__init__.py -> build/lib/apex/RNN
    copying apex/RNN/RNNBackend.py -> build/lib/apex/RNN
    copying apex/RNN/models.py -> build/lib/apex/RNN
    creating build/lib/apex/contrib
    copying apex/contrib/__init__.py -> build/lib/apex/contrib
    creating build/lib/apex/amp/lists
    copying apex/amp/lists/__init__.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/tensor_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/torch_overrides.py -> build/lib/apex/amp/lists
    copying apex/amp/lists/functional_overrides.py -> build/lib/apex/amp/lists
    creating build/lib/apex/transformer/amp
    copying apex/transformer/amp/__init__.py -> build/lib/apex/transformer/amp
    copying apex/transformer/amp/grad_scaler.py -> build/lib/apex/transformer/amp
    creating build/lib/apex/transformer/_data
    copying apex/transformer/_data/__init__.py -> build/lib/apex/transformer/_data
    copying apex/transformer/_data/_batch_sampler.py ->

```

```

build/lib/apex/transformer/_data
    creating build/lib/apex/transformer/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/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
    copying apex/transformer/pipeline_parallel/p2p_communication.py ->
build/lib/apex/transformer/pipeline_parallel
    copying apex/transformer/pipeline_parallel/_timers.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/tensor_parallel
    copying apex/transformer/tensor_parallel/memory.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/utils.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/__init__.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/mappings.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/data.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/layers.py ->
build/lib/apex/transformer/tensor_parallel
    copying apex/transformer/tensor_parallel/random.py ->
build/lib/apex/transformer/tensor_parallel
    creating build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_gpt.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/__init__.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_bert.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/distributed_test_base.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/arguments.py ->
build/lib/apex/transformer/testing

```

```

    copying apex/transformer/testing/commons.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/global_vars.py ->
build/lib/apex/transformer/testing
    copying apex/transformer/testing/standalone_transformer_lm.py ->
build/lib/apex/transformer/testing
    creating 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
    copying apex/transformer/pipeline_parallel/schedules/common.py ->
build/lib/apex/transformer/pipeline_parallel/schedules
    creating build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/__init__.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck_module_test.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/halo_exchangers.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/bottleneck.py ->
build/lib/apex/contrib/bottleneck
    copying apex/contrib/bottleneck/test.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/groupbn
    copying apex/contrib/groupbn/__init__.py -> build/lib/apex/contrib/groupbn
    copying apex/contrib/groupbn/batch_norm.py -> build/lib/apex/contrib/groupbn
    creating build/lib/apex/contrib/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_self_multihead_attn_func.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/fast_encdec_multihead_attn_norm_add_func.py ->

```

```

build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/self_multihead_attn.py ->
build/lib/apex/contrib/multihead_attn
    copying apex/contrib/multihead_attn/mask_softmax_dropout_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/self_multihead_attn_func.py ->
build/lib/apex/contrib/multihead_attn
    creating build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/__init__.py ->
build/lib/apex/contrib/index_mul_2d
    copying apex/contrib/index_mul_2d/index_mul_2d.py ->
build/lib/apex/contrib/index_mul_2d
    creating build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fp16_optimizer.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_sgd.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/__init__.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_adam.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/distributed_fused_lamb.py ->
build/lib/apex/contrib/optimizers
    copying apex/contrib/optimizers/fused_lamb.py ->
build/lib/apex/contrib/optimizers
    creating build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/__init__.py -> build/lib/apex/contrib/fmha
    copying apex/contrib/fmha/fmha.py -> build/lib/apex/contrib/fmha
    creating build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/__init__.py -> build/lib/apex/contrib/xentropy
    copying apex/contrib/xentropy/softmax_xentropy.py ->
build/lib/apex/contrib/xentropy
    creating build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/permutation_lib.py ->
build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/__init__.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/asp.py -> build/lib/apex/contrib/sparsity
    copying apex/contrib/sparsity/sparse_masklib.py ->
build/lib/apex/contrib/sparsity
    creating build/lib/apex/contrib/layer_norm
    copying apex/contrib/layer_norm/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/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_halo_exchanger_1d.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/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/clip_grad
    copying apex/contrib/clip_grad/__init__.py -> build/lib/apex/contrib/clip_grad
    copying apex/contrib/clip_grad/clip_grad.py ->
build/lib/apex/contrib/clip_grad
    creating build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/__init__.py ->
build/lib/apex/contrib/focal_loss
    copying apex/contrib/focal_loss/focal_loss.py ->
build/lib/apex/contrib/focal_loss
    creating build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
-> build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/__init__.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying
apex/contrib/sparsity/permutation_search_kernels/permutation_utilities.py ->
build/lib/apex/contrib/sparsity/permutation_search_kernels
    copying apex/contrib/sparsity/permutation_search_kernels/call_permutation_sear
ch_kernels.py -> build/lib/apex/contrib/sparsity/permutation_search_kernels
    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/normalization
    copying build/lib/apex/normalization/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    copying build/lib/apex/normalization/fused_layer_norm.py ->
build/bdist.linux-x86_64/wheel/apex/normalization
    creating build/bdist.linux-x86_64/wheel/apex/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/amp
    copying build/lib/apex/amp/opt.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/scaler.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/frontend.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/__version__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/utils.py -> build/bdist.linux-x86_64/wheel/apex/amp
    copying build/lib/apex/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp
    creating build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/tensor_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/torch_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/lists/functional_overrides.py ->
build/bdist.linux-x86_64/wheel/apex/amp/lists
    copying build/lib/apex/amp/amp.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/wrap.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/_process_optimizer.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/__init__.py -> build/bdist.linux-x86_64/wheel/apex
    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/fused_dense
    copying build/lib/apex/fused_dense/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    copying build/lib/apex/fused_dense/fused_dense.py ->
build/bdist.linux-x86_64/wheel/apex/fused_dense
    creating build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/parallel

```

```

    copying build/lib/apex/parallel/optimized_sync_batchnorm_kernel.py ->
build/bdist.linux-x86_64/wheel/apex/parallel
    copying build/lib/apex/parallel/multiproc.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
    copying build/lib/apex/parallel/distributed.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
    creating 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_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
    copying build/lib/apex/optimizers/fused_adagrad.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_novograd.py ->
build/bdist.linux-x86_64/wheel/apex/optimizers
    copying build/lib/apex/optimizers/fused_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/fp16_optimizer.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/fp16util.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    copying build/lib/apex/fp16_utils/loss_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/fp16_utils
    creating build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/amp/grad_scaler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/amp
    copying build/lib/apex/transformer/microbatches.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/__init__.py ->

```



```

build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/_data/_batch_sampler.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/_data
    copying build/lib/apex/transformer/parallel_state.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    copying build/lib/apex/transformer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer
    creating build/bdist.linux-x86_64/wheel/apex/transformer/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
    creating 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
    copying build/lib/apex/transformer/pipeline_parallel/p2p_communication.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    creating
build/bdist.linux-x86_64/wheel/apex/transformer/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/schedules/common.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel/schedules
    copying build/lib/apex/transformer/pipeline_parallel/_timers.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/pipeline_parallel
    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/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/memory.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/utils.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel

```

```

    copying build/lib/apex/transformer/tensor_parallel/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/mappings.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/tensor_parallel
    copying build/lib/apex/transformer/tensor_parallel/data.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/layers.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
    creating 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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_bert.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/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/testing/commons.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/global_vars.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    copying build/lib/apex/transformer/testing/standalone_transformer_lm.py ->
build/bdist.linux-x86_64/wheel/apex/transformer/testing
    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/RNN
    copying build/lib/apex/RNN/cells.py -> build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/RNNBackend.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    copying build/lib/apex/RNN/models.py ->
build/bdist.linux-x86_64/wheel/apex/RNN
    creating build/bdist.linux-x86_64/wheel/apex/contrib
    creating build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/bottleneck_module_test.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
    copying build/lib/apex/contrib/bottleneck/halo_exchangers.py ->

```

```

build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/bottleneck.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/bottleneck
  copying build/lib/apex/contrib/bottleneck/test.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
  copying build/lib/apex/contrib/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib
  creating build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  copying build/lib/apex/contrib/groupbn/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  copying build/lib/apex/contrib/groupbn/batch_norm.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/groupbn
  creating build/bdist.linux-x86_64/wheel/apex/contrib/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_self_multihead_attn_func.py
-> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying
build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/fast_encdec_multihead_attn_norm_
add_func.py -> build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/mask_softmax_dropout_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/encdec_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  copying build/lib/apex/contrib/multihead_attn/self_multihead_attn_func.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/multihead_attn
  creating build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  copying build/lib/apex/contrib/index_mul_2d/index_mul_2d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/index_mul_2d
  creating build/bdist.linux-x86_64/wheel/apex/contrib/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_sgd.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_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_adam.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/optimizers
    copying build/lib/apex/contrib/optimizers/distributed_fused_lamb.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
    creating build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    copying build/lib/apex/contrib/fmha/fmha.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/fmha
    creating build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    copying build/lib/apex/contrib/xentropy/softmax_xentropy.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/xentropy
    creating build/bdist.linux-x86_64/wheel/apex/contrib/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/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    copying build/lib/apex/contrib/sparsity/asp.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    creating
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying
build/lib/apex/contrib/sparsity/permutation_search_kernels/exhaustive_search.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/__init__.py
->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/permutation
_utilities.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/permutation_search_kernels/call_permut
ation_search_kernels.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity/permutation_search_kernels
    copying build/lib/apex/contrib/sparsity/sparse_masklib.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/sparsity
    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/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_halo_exchanger_1d.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  copying build/lib/apex/contrib/peer_memory/peer_memory.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/peer_memory
  creating build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  copying build/lib/apex/contrib/transducer/transducer.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/transducer
  creating build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  copying build/lib/apex/contrib/clip_grad/clip_grad.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/clip_grad
  creating build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/__init__.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/contrib/focal_loss/focal_loss.py ->
build/bdist.linux-x86_64/wheel/apex/contrib/focal_loss
  copying build/lib/apex/_autocast_utils.py ->
build/bdist.linux-x86_64/wheel/apex
  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-8lq2_k3d/apex-0.1-py3-none-any.whl' and adding
'build/bdist.linux-x86_64/wheel' to it
  adding 'apex/__init__.py'
  adding 'apex/_autocast_utils.py'
  adding 'apex/RNN/RNNBackend.py'

```

```

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

```

```

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

```

```

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

```



```

adding 'apex/transformer/testing/standalone_transformer_lm.py'
adding 'apex-0.1.dist-info/LICENSE'
adding 'apex-0.1.dist-info/METADATA'
adding 'apex-0.1.dist-info/WHEEL'
adding 'apex-0.1.dist-info/top_level.txt'
adding 'apex-0.1.dist-info/RECORD'
removing build/bdist.linux-x86_64/wheel
done
Created wheel for apex: filename=apex-0.1-py3-none-any.whl size=298114
sha256=34209e53ca7fb8d25b87901d4419c1ebd795b4f8194c5f90209f35e31ba9f7e2
Stored in directory: /tmp/pip-ephem-wheel-cache-
vvkbei0i/wheels/6b/22/a2/1665526ee3c3061243260fd522525eca31398e04bfa5ad7e1f
Successfully built apex
Installing collected packages: apex
Value for scheme.platlib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.purelib does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/lib/python3.7/dist-packages
sysconfig: /usr/lib/python3.7/site-packages
Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/apex
sysconfig: /usr/include/python3.7m/apex
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None

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

```

```

Value for scheme.headers does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/include/python3.7/UNKNOWN
sysconfig: /usr/include/python3.7m/UNKNOWN
Value for scheme.scripts does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local/bin
sysconfig: /usr/bin
Value for scheme.data does not match. Please report this to
<https://github.com/pypa/pip/issues/9617>
distutils: /usr/local
sysconfig: /usr
Additional context:
user = False
home = None
root = None
prefix = None
Successfully installed apex-0.1
Removed build tracker: '/tmp/pip-req-tracker-c2evqbtv'
/content/ditto

```

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting urllib3==1.25.4
  Downloading urllib3-1.25.4-py2.py3-none-any.whl (125 kB)
    |                               | 125 kB 4.9 MB/s
Collecting awscli
  Downloading awscli-1.25.47-py3-none-any.whl (3.9 MB)
    |                               | 3.9 MB 40.5 MB/s
Collecting botocore==1.27.47
  Downloading botocore-1.27.47-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 8.8 MB/s
Collecting docutils<0.17,>=0.10
  Downloading docutils-0.16-py2.py3-none-any.whl (548 kB)
    |                               | 548 kB 14.0 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)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in

```

```

/usr/local/lib/python3.7/dist-packages (from botocore==1.27.47->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==1.27.47->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-1.25.47 botocore-1.27.47 colorama-0.4.4
docutils-0.16 jmespath-1.0.1 rsa-4.7.2 s3transfer-0.6.0 urllib3-1.25.4

```

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

```

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

```

```

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

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

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

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

```

0.3 importa config.json

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

```

Cloning into 'IC'...
remote: Enumerating objects: 982, done.
remote: Counting objects: 100% (52/52), done.
remote: Compressing objects: 100% (48/48), done.
remote: Total 982 (delta 16), reused 7 (delta 0), pack-reused 930
Receiving objects: 100% (982/982), 251.89 MiB | 28.28 MiB/s, done.
Resolving deltas: 100% (600/600), done.
Checking out files: 100% (1273/1273), done.

```

0.4 remove e move

```
[ ]: !ls
```

```

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

```

```
[ ]: !rm configs.json
!mv -n /content/ditto/IC/datasasErros/BDirty_50_3_5/configs.json /content/ditto/
```

1 DA

1.1 DA STRUCTURED

1.1.1 Beer - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm 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, 429kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 2.11MB/s]

Downloading: 100% 440M/440M [00:09<00:00, 48.2MB/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: 0.5753392577171326
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.3170731707317073, f1=0.37681159420289856,
best_f1=0.37681159420289856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.496731698513031
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

```

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=0.41666666666666663, f1=0.3333333333333333,
best_f1=0.3333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5465734004974365
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5128205128205129, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.30273374915122986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5641025641025641, 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: 0.3335566818714142
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5853658536585367, f1=0.4444444444444444,
best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2905455529689789

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7058823529411764, f1=0.45161290322580644,
best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22130781412124634
/usr/local/lib/python3.7/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=0.45454545454545453, f1=0.47999999999999999,
best_f1=0.45161290322580644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08355729281902313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.75, f1=0.4827586206896552, best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.031443554908037186
/usr/local/lib/python3.7/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=0.7096774193548386, f1=0.5517241379310344,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.037591997534036636
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7878787878787878, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014551480300724506
/usr/local/lib/python3.7/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=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023498041555285454
/usr/local/lib/python3.7/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=0.7878787878787878, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04131972789764404
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8125000000000001, f1=0.5333333333333333,
best_f1=0.5333333333333333
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015853818506002426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8387096774193549, f1=0.5333333333333333,
best_f1=0.5333333333333333

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009759517386555672
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.8387096774193549, f1=0.5333333333333333,
best_f1=0.5333333333333333
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 107606.90it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7333333333333334
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 [00:15, 285.85it/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: 0.6013573408126831
step: 10, loss: 0.6070559620857239
step: 20, loss: 0.3360409736633301
step: 30, loss: 0.17598143219947815
step: 40, loss: 0.27434656023979187
step: 50, loss: 0.026675183326005936
step: 60, loss: 0.06487015634775162
step: 70, loss: 0.08455372601747513
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 80, loss: 0.08345964550971985
step: 90, loss: 0.0673287957906723
step: 100, loss: 0.012675479985773563
step: 110, loss: 0.033842090517282486
step: 120, loss: 0.006999011617153883
step: 130, loss: 0.009491421282291412
step: 140, loss: 0.0035409170668572187
step: 150, loss: 0.0402308814227581
step: 160, loss: 0.011047862470149994
step: 170, loss: 0.0774821862578392
step: 180, loss: 0.1287127584218979
step: 190, loss: 0.022490086033940315
step: 200, loss: 0.013345815241336823
step: 210, loss: 0.005647131707519293
step: 220, loss: 0.011497896164655685
step: 230, loss: 0.07803533226251602
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9797297297297298, 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: 0.008022038266062737
step: 10, loss: 0.003938931040465832
step: 20, loss: 0.07578703761100769
step: 30, loss: 0.19600339233875275
step: 40, loss: 0.010931318625807762
step: 50, loss: 0.011455580592155457
step: 60, loss: 0.003729761578142643
step: 70, loss: 0.16403187811374664
step: 80, loss: 0.0024412444327026606
step: 90, loss: 0.04198402911424637
step: 100, loss: 0.01045041810721159
step: 110, loss: 0.05084560811519623
step: 120, loss: 0.008762889541685581
step: 130, loss: 0.01731971651315689
step: 140, loss: 0.0020206684712320566
step: 150, loss: 0.04161320626735687
step: 160, loss: 0.08423975855112076
step: 170, loss: 0.006476460490375757
step: 180, loss: 0.009269333444535732
step: 190, loss: 0.006837426219135523
step: 200, loss: 0.00602988013997674
step: 210, loss: 0.0026552702765911818
step: 220, loss: 0.09536083042621613
step: 230, loss: 0.013209945522248745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9853768278965129, f1=0.9820224719101124,
best_f1=0.9820224719101124
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030039316043257713

```

```

step: 10, loss: 0.015411989763379097
step: 20, loss: 0.0005327625549398363
step: 30, loss: 0.07177217304706573
step: 40, loss: 0.19674301147460938
step: 50, loss: 0.011488345451653004
step: 60, loss: 0.053095266222953796
step: 70, loss: 0.009720377624034882
step: 80, loss: 0.0011898879893124104
step: 90, loss: 0.027171041816473007
step: 100, loss: 0.0041281781159341335
step: 110, loss: 0.013434148393571377
step: 120, loss: 0.021597618237137794
step: 130, loss: 0.001095907879061997
step: 140, loss: 0.023736651986837387
step: 150, loss: 0.0023764583747833967
step: 160, loss: 0.007658728398382664
step: 170, loss: 0.003205052576959133
step: 180, loss: 0.004963978659361601
step: 190, loss: 0.0026105064898729324
step: 200, loss: 0.007440219633281231
step: 210, loss: 0.0007694665109738708
step: 220, loss: 0.0007738024578429759
step: 230, loss: 0.00028531765565276146
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9721913236929923,
best_f1=0.9820224719101124
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021554611157625914
step: 10, loss: 0.0010465357918292284
step: 20, loss: 0.0007259598933160305
step: 30, loss: 0.0007204951252788305
step: 40, loss: 0.002488137222826481
step: 50, loss: 0.0006951417890377343
step: 60, loss: 0.004837350454181433
step: 70, loss: 0.006144417449831963
step: 80, loss: 0.0032744244672358036
step: 90, loss: 0.021200409159064293
step: 100, loss: 0.0006606568349525332
step: 110, loss: 0.0006986315129324794
step: 120, loss: 0.006224417593330145
step: 130, loss: 0.018937446177005768
step: 140, loss: 0.0024467853363603354
step: 150, loss: 0.1960906833410263
step: 160, loss: 0.0010241226991638541

```

```

step: 170, loss: 0.11325991898775101
step: 180, loss: 0.0003567279491107911
step: 190, loss: 0.001407412812113762
step: 200, loss: 0.007644502446055412
step: 210, loss: 0.09869041293859482
step: 220, loss: 0.0003465230984147638
step: 230, loss: 0.017895342782139778
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9864253393665158, f1=0.978675645342312,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024687605910003185
step: 10, loss: 0.012674933299422264
step: 20, loss: 0.00432887626811862
step: 30, loss: 0.005247761495411396
step: 40, loss: 0.00037246374995447695
step: 50, loss: 0.0003921421302948147
step: 60, loss: 0.01886470802128315
step: 70, loss: 0.0013022724306210876
step: 80, loss: 0.0003783321881201118
step: 90, loss: 0.0005042756674811244
step: 100, loss: 0.0008195051923394203
step: 110, loss: 0.00044350652024149895
step: 120, loss: 0.00011885316052939743
step: 130, loss: 0.006516129709780216
step: 140, loss: 0.0005343113443814218
step: 150, loss: 0.00033529510255903006
step: 160, loss: 0.03897297382354736
step: 170, loss: 0.08714138716459274
step: 180, loss: 0.018664497882127762
step: 190, loss: 0.006379565689712763
step: 200, loss: 0.0022215035278350115
step: 210, loss: 0.000739425653591752
step: 220, loss: 0.00039744647801853716
step: 230, loss: 0.00033710492425598204
/usr/local/lib/python3.7/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=0.9795454545454545, f1=0.9775280898876404,
best_f1=0.978675645342312

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038718702853657305
step: 10, loss: 0.0002513524377718568
step: 20, loss: 0.00596547219902277
step: 30, loss: 0.0002999426214955747
step: 40, loss: 0.006828164216130972
step: 50, loss: 0.0036844939459115267
step: 60, loss: 0.03384861722588539
step: 70, loss: 0.009600004181265831
step: 80, loss: 0.0008418120560236275
step: 90, loss: 0.0005064926808699965
step: 100, loss: 0.001678977394476533
step: 110, loss: 0.00019402675388846546
step: 120, loss: 0.00039947914774529636
step: 130, loss: 0.000692762085236609
step: 140, loss: 0.0001351213431917131
step: 150, loss: 0.0002154184621758759
step: 160, loss: 0.0020444272086024284
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 170, loss: 0.05686277896165848
step: 180, loss: 0.00512382248416543
step: 190, loss: 0.0831354409456253
step: 200, loss: 0.00013412110274657607
step: 210, loss: 0.00058198586339131
step: 220, loss: 0.0008083231514319777
step: 230, loss: 0.10388009250164032
/usr/local/lib/python3.7/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=0.9853438556933484, f1=0.9843400447427293,
best_f1=0.978675645342312
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016903647920116782
step: 10, loss: 0.00032057930366136134
step: 20, loss: 0.002689382992684841
step: 30, loss: 0.0005830821464769542
step: 40, loss: 0.00020939725800417364
step: 50, loss: 0.00040131848072633147
step: 60, loss: 0.038178540766239166
step: 70, loss: 0.0008863491821102798
step: 80, loss: 0.0009946455247700214
step: 90, loss: 0.0004078477213624865
step: 100, loss: 0.001387831405736506
step: 110, loss: 0.0006794329383410513

```



```

step: 120, loss: 0.0007332301465794444
step: 130, loss: 0.000375247560441494
step: 140, loss: 0.0003962013288401067
step: 150, loss: 0.00028758178814314306
step: 160, loss: 0.012042982503771782
step: 170, loss: 0.009263628162443638
step: 180, loss: 0.0005502121639437973
step: 190, loss: 0.00019766493642237037
step: 200, loss: 0.009019981138408184
step: 210, loss: 6.882608431624249e-05
step: 220, loss: 0.00034216701169498265
step: 230, loss: 0.0001804190978873521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9875706214689265, f1=0.9776286353467561,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001070320577127859
step: 10, loss: 0.0005086367018520832
step: 20, loss: 0.0002758812333922833
step: 30, loss: 0.0003331276820972562
step: 40, loss: 0.0005960497073829174
step: 50, loss: 0.00011021069076377898
step: 60, loss: 7.490520511055365e-05
step: 70, loss: 0.00010600591485854238
step: 80, loss: 0.00030932563822716475
step: 90, loss: 6.276799831539392e-05
step: 100, loss: 0.00019276607781648636
step: 110, loss: 9.94531437754631e-05
step: 120, loss: 0.0008630756055936217
step: 130, loss: 6.739694799762219e-05
step: 140, loss: 5.6998844229383394e-05
step: 150, loss: 0.0001310209627263248
step: 160, loss: 0.000862397369928658
step: 170, loss: 0.00020116080122534186
step: 180, loss: 0.0005355095490813255
step: 190, loss: 0.00010895109153352678
step: 200, loss: 0.0001543721737107262
step: 210, loss: 0.0002128013875335455
step: 220, loss: 0.0006404372397810221
step: 230, loss: 0.00020229702931828797

```

```

/usr/local/lib/python3.7/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=0.9842342342342343, f1=0.9755555555555556,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00017261509492527694
step: 10, loss: 0.00047462049406021833
step: 20, loss: 0.0005435242783278227
step: 30, loss: 6.594535807380453e-05
step: 40, loss: 0.032712288200855255
step: 50, loss: 0.000188158723176457
step: 60, loss: 0.0001001627606456168
step: 70, loss: 0.00010224637662759051
step: 80, loss: 0.02448841743171215
step: 90, loss: 0.004334176890552044
step: 100, loss: 0.000435000954894349
step: 110, loss: 5.171049997443333e-05
step: 120, loss: 6.940917228348553e-05
step: 130, loss: 7.11390093783848e-05
step: 140, loss: 4.8936035454971716e-05
step: 150, loss: 0.002370311878621578
step: 160, loss: 0.0006379553815349936
step: 170, loss: 0.00015498597349505872
step: 180, loss: 0.12832209467887878
step: 190, loss: 0.00646374374628067
step: 200, loss: 0.0005295358132570982
step: 210, loss: 0.0005585363251157105
step: 220, loss: 0.00012665135727729648
step: 230, loss: 0.00012818945106118917
/usr/local/lib/python3.7/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=0.9830124575311437, f1=0.9809203142536477,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013583940744865686
step: 10, loss: 5.4440577514469624e-05
step: 20, loss: 0.00012671096192207187
step: 30, loss: 0.0011073241475969553
step: 40, loss: 5.93911754549481e-05
step: 50, loss: 0.011771738529205322
step: 60, loss: 0.0005578835844062269
step: 70, loss: 8.452167821815237e-05

```

```

step: 80, loss: 9.731640602694824e-05
step: 90, loss: 0.00015772362530697137
step: 100, loss: 9.106705692829564e-05
step: 110, loss: 0.0001118216387112625
step: 120, loss: 0.0012280871160328388
step: 130, loss: 5.7776029279921204e-05
step: 140, loss: 0.00011002847168128937
step: 150, loss: 0.00014525694132316858
step: 160, loss: 2.611711352074053e-05
step: 170, loss: 3.7444444256881252e-05
step: 180, loss: 0.00011745860683731735
step: 190, loss: 0.00013000897888559848
step: 200, loss: 4.6838449634378776e-05
step: 210, loss: 0.0007462245994247496
step: 220, loss: 0.017785081639885902
step: 230, loss: 0.00047537742648273706
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9820627802690582,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000121133285574615
step: 10, loss: 0.0001619068207219243
step: 20, loss: 0.0001906638644868508
step: 30, loss: 0.0001535180490463972
step: 40, loss: 0.004485145211219788
step: 50, loss: 0.00010638878302415833
step: 60, loss: 0.0001314105320489034
step: 70, loss: 0.00044077468919567764
step: 80, loss: 4.915962563245557e-05
step: 90, loss: 0.00013561194646172225
step: 100, loss: 0.00014885970449540764
step: 110, loss: 5.990053978166543e-05
step: 120, loss: 8.93393371370621e-05
step: 130, loss: 0.0003833512600976974
step: 140, loss: 0.00012178211909485981
step: 150, loss: 0.002321881242096424
step: 160, loss: 4.21497825300321e-05
step: 170, loss: 7.341026503127068e-05
step: 180, loss: 9.771653276402503e-05
step: 190, loss: 5.47205563634634e-05
step: 200, loss: 8.378973143408075e-05
step: 210, loss: 0.00010053018922917545
step: 220, loss: 5.1182640163460746e-05
step: 230, loss: 0.00016221955593209714

```

```

/usr/local/lib/python3.7/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=0.9853107344632768, f1=0.9787234042553192,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.640360566554591e-05
step: 10, loss: 0.000158798458869569
step: 20, loss: 7.261527935042977e-05
step: 30, loss: 0.0029541479889303446
step: 40, loss: 6.601745553780347e-05
step: 50, loss: 7.890408596722409e-05
step: 60, loss: 0.00010680993727874011
step: 70, loss: 0.0002971779613289982
step: 80, loss: 0.0002940516860689968
step: 90, loss: 4.9379854317521676e-05
step: 100, loss: 5.428829535958357e-05
step: 110, loss: 0.00010257303802063689
step: 120, loss: 0.0001378476881654933
step: 130, loss: 4.744710531667806e-05
step: 140, loss: 8.728313696337864e-05
step: 150, loss: 0.00021239681518636644
step: 160, loss: 4.885893577011302e-05
step: 170, loss: 9.024206519825384e-05
step: 180, loss: 0.00010667569586075842
step: 190, loss: 6.770806066924706e-05
step: 200, loss: 5.520151535165496e-05
step: 210, loss: 9.356454393127933e-05
step: 220, loss: 6.662544910795987e-05
step: 230, loss: 0.045218296349048615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9853768278965129, f1=0.9777777777777777,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003606655518524349
step: 10, loss: 0.0011080448748543859
step: 20, loss: 0.0004454113950487226
step: 30, loss: 8.668890222907066e-05
step: 40, loss: 0.0007098490023054183
step: 50, loss: 5.148863419890404e-05
step: 60, loss: 5.986414180370048e-05
step: 70, loss: 0.001369922305457294

```

```

step: 80, loss: 5.07326731167268e-05
step: 90, loss: 0.09111092984676361
step: 100, loss: 2.7812224288936704e-05
step: 110, loss: 0.0003105713112745434
step: 120, loss: 0.0002807806304190308
step: 130, loss: 0.0019387674983590841
step: 140, loss: 3.849569839076139e-05
step: 150, loss: 6.316399230854586e-05
step: 160, loss: 3.4305048757232726e-05
step: 170, loss: 0.0001381261390633881
step: 180, loss: 6.75842966302298e-05
step: 190, loss: 3.495247074170038e-05
step: 200, loss: 0.00012198969488963485
step: 210, loss: 2.8329912311164662e-05
step: 220, loss: 0.0002327479305677116
step: 230, loss: 3.962752452935092e-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=0.9841269841269841, f1=0.9831271091113611,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.0682217988651246e-05
step: 10, loss: 7.019712938927114e-05
step: 20, loss: 8.032324694795534e-05
step: 30, loss: 5.743187284679152e-05
step: 40, loss: 5.324011362972669e-05
step: 50, loss: 4.333762262831442e-05
step: 60, loss: 3.780248880502768e-05
step: 70, loss: 6.109176320023835e-05
step: 80, loss: 0.0001799958699848503
step: 90, loss: 6.49614303256385e-05
step: 100, loss: 5.043841156293638e-05
step: 110, loss: 3.473660399322398e-05
step: 120, loss: 5.838511424371973e-05
step: 130, loss: 6.552097329404205e-05
step: 140, loss: 8.031503239180893e-05
step: 150, loss: 6.388021574821323e-05
step: 160, loss: 0.0008827882120385766
step: 170, loss: 3.395454405108467e-05
step: 180, loss: 4.47346901637502e-05
step: 190, loss: 5.025597783969715e-05
step: 200, loss: 4.732410525321029e-05
step: 210, loss: 5.788373528048396e-05
step: 220, loss: 5.7231998653151095e-05
step: 230, loss: 5.851074820384383e-05

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9853768278965129, f1=0.9799107142857142,
best_f1=0.9776286353467561
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.1444989619776607e-05
step: 10, loss: 2.4441162167931907e-05
step: 20, loss: 0.0001459898048778996
step: 30, loss: 4.6522683987859637e-05
step: 40, loss: 0.0007660037372261286
step: 50, loss: 0.010284139774739742
step: 60, loss: 4.3758656829595566e-05
step: 70, loss: 4.199723116471432e-05
step: 80, loss: 0.00013999160728417337
step: 90, loss: 0.003579790471121669
step: 100, loss: 0.0005790345603600144
step: 110, loss: 6.433464295696467e-05
step: 120, loss: 6.402860162779689e-05
step: 130, loss: 0.0001414270664099604
step: 140, loss: 3.345579534652643e-05
step: 150, loss: 0.0001055607499438338
step: 160, loss: 3.618956543505192e-05
step: 170, loss: 7.966989505803213e-05
step: 180, loss: 8.642025204608217e-05
step: 190, loss: 9.063282777788118e-05
step: 200, loss: 0.00010273526277160272
step: 210, loss: 4.5516484533436596e-05
step: 220, loss: 0.000106295759906061
step: 230, loss: 5.0993243348784745e-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=0.9841986455981941, f1=0.980963045912654,
best_f1=0.9776286353467561
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, 219.66it/s]
load_f1 = 0.9864559819413092
real_f1 = 0.9875706214689265
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 267.93it/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, 301kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 254kB/s]

Downloading: 100% 440M/440M [00:07<00:00, 55.2MB/s]

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

```
enabled          : True
opt_level        : 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: 0.6292595267295837
step: 10, loss: 0.5241456627845764
step: 20, loss: 0.5412311553955078
step: 30, loss: 0.2758166491985321
step: 40, loss: 0.21162381768226624
step: 50, loss: 0.2294643521308899
step: 60, loss: 0.07686296105384827
step: 70, loss: 0.1794697493314743
step: 80, loss: 0.06249404698610306
step: 90, loss: 0.23383788764476776
step: 100, loss: 0.03747272118926048
step: 110, loss: 0.16519811749458313
step: 120, loss: 0.1257106512784958
step: 130, loss: 0.03817928954958916
step: 140, loss: 0.07654178887605667
step: 150, loss: 0.1262238621711731
step: 160, loss: 0.055399730801582336
step: 170, loss: 0.2391255497932434
step: 180, loss: 0.11300566047430038
step: 190, loss: 0.018327001482248306
step: 200, loss: 0.14910788834095
step: 210, loss: 0.06383755803108215
step: 220, loss: 0.2640988230705261
step: 230, loss: 0.2049870640039444
step: 240, loss: 0.08612660318613052
step: 250, loss: 0.04586506634950638
step: 260, loss: 0.15388384461402893
step: 270, loss: 0.012347766198217869
step: 280, loss: 0.06911744922399521
step: 290, loss: 0.23481285572052002
step: 300, loss: 0.1338557004928589
step: 310, loss: 0.18356655538082123
step: 320, loss: 0.06826626509428024
step: 330, loss: 0.052743468433618546
step: 340, loss: 0.0949038565158844
step: 350, loss: 0.06264176219701767
step: 360, loss: 0.11687801033258438
step: 370, loss: 0.128071129322052
step: 380, loss: 0.02153776027262211
step: 390, loss: 0.1777782291173935
step: 400, loss: 0.28476157784461975
step: 410, loss: 0.07342232018709183
```

```

step: 420, loss: 0.09527470171451569
step: 430, loss: 0.2101748138666153
step: 440, loss: 0.048064280301332474
step: 450, loss: 0.02241206169128418
step: 460, loss: 0.08090080320835114
step: 470, loss: 0.11579643189907074
step: 480, loss: 0.12319420278072357
step: 490, loss: 0.08605089783668518
step: 500, loss: 0.09878551214933395
step: 510, loss: 0.031415410339832306
step: 520, loss: 0.03261495381593704
step: 530, loss: 0.006239640060812235
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9177126917712692, f1=0.9094287041337669,
best_f1=0.9094287041337669
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1855548769235611
step: 10, loss: 0.09075064212083817
step: 20, loss: 0.05159248411655426
step: 30, loss: 0.015181455761194229
step: 40, loss: 0.06025256961584091
step: 50, loss: 0.12272150069475174
step: 60, loss: 0.022403771057724953
step: 70, loss: 0.04784879833459854
step: 80, loss: 0.039197538048028946
step: 90, loss: 0.020221777260303497
step: 100, loss: 0.018827252089977264
step: 110, loss: 0.07474356144666672
step: 120, loss: 0.08009224385023117
step: 130, loss: 0.11337782442569733
step: 140, loss: 0.06185245141386986
step: 150, loss: 0.0706922858953476
step: 160, loss: 0.01749420538544655
step: 170, loss: 0.02778957411646843
step: 180, loss: 0.04799700155854225
step: 190, loss: 0.06773550808429718
step: 200, loss: 0.025821393355727196
step: 210, loss: 0.06775090843439102
step: 220, loss: 0.05011194571852684
step: 230, loss: 0.0023293967824429274

```

```

step: 240, loss: 0.1620742231607437
step: 250, loss: 0.038859229534864426
step: 260, loss: 0.006959986872971058
step: 270, loss: 0.2919214963912964
step: 280, loss: 0.010530252940952778
step: 290, loss: 0.06518350541591644
step: 300, loss: 0.17350277304649353
step: 310, loss: 0.018939483910799026
step: 320, loss: 0.08419772237539291
step: 330, loss: 0.11764539033174515
step: 340, loss: 0.023456400260329247
step: 350, loss: 0.003406904172152281
step: 360, loss: 0.07894114404916763
step: 370, loss: 0.1006658673286438
step: 380, loss: 0.04969139024615288
step: 390, loss: 0.06612755358219147
step: 400, loss: 0.1343858391046524
step: 410, loss: 0.04157615080475807
step: 420, loss: 0.05197615921497345
step: 430, loss: 0.011977686546742916
step: 440, loss: 0.023150956258177757
step: 450, loss: 0.005144699010998011
step: 460, loss: 0.061785031110048294
step: 470, loss: 0.0518442802131176
step: 480, loss: 0.3840582072734833
step: 490, loss: 0.016895651817321777
step: 500, loss: 0.38235536217689514
step: 510, loss: 0.02842988818883896
step: 520, loss: 0.06211164593696594
step: 530, loss: 0.07814369350671768
/usr/local/lib/python3.7/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=0.9172445677300047, f1=0.9143914854234151,
best_f1=0.9094287041337669
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20274955034255981
step: 10, loss: 0.11166493594646454
step: 20, loss: 0.06103145331144333
step: 30, loss: 0.20120938122272491
step: 40, loss: 0.023757854476571083
step: 50, loss: 0.06861714273691177
step: 60, loss: 0.010245643556118011
step: 70, loss: 0.0036424368154257536
step: 80, loss: 0.00039130469667725265
step: 90, loss: 0.0031815001275390387

```

```
step: 100, loss: 0.0479995422065258
step: 110, loss: 0.010031557641923428
step: 120, loss: 0.007067587226629257
step: 130, loss: 0.0019744131714105606
step: 140, loss: 0.014156432822346687
step: 150, loss: 0.022721366956830025
step: 160, loss: 0.007276264484971762
step: 170, loss: 0.050316378474235535
step: 180, loss: 0.04257262870669365
step: 190, loss: 0.0432916060090065
step: 200, loss: 0.11889093369245529
step: 210, loss: 0.027320140972733498
step: 220, loss: 0.16965264081954956
step: 230, loss: 0.16168726980686188
step: 240, loss: 0.016871383413672447
step: 250, loss: 0.017369598150253296
step: 260, loss: 0.022336456924676895
step: 270, loss: 0.01240269560366869
step: 280, loss: 0.09822557866573334
step: 290, loss: 0.009159758687019348
step: 300, loss: 0.03079536370933056
step: 310, loss: 0.004888387396931648
step: 320, loss: 0.033161673694849014
step: 330, loss: 0.0014068839373067021
step: 340, loss: 0.039964620023965836
step: 350, loss: 0.08842281997203827
step: 360, loss: 0.07358375936746597
step: 370, loss: 0.028494877740740776
step: 380, loss: 0.03684544563293457
step: 390, loss: 0.008694295771420002
step: 400, loss: 0.008539127185940742
step: 410, loss: 0.020855195820331573
step: 420, loss: 0.1981791853904724
step: 430, loss: 0.005431808531284332
step: 440, loss: 0.005484815686941147
step: 450, loss: 0.08768568187952042
step: 460, loss: 0.05184788256883621
step: 470, loss: 0.07201480120420456
step: 480, loss: 0.017078375443816185
step: 490, loss: 0.005058600567281246
step: 500, loss: 0.008187989704310894
step: 510, loss: 0.0064084772020578384
step: 520, loss: 0.017921103164553642
step: 530, loss: 0.07256343960762024
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
```

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=0.9257884972170687, f1=0.9136490250696379,

best_f1=0.9136490250696379

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

An input tensor was not cuda.

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

step: 0, loss: 0.012251448817551136

step: 10, loss: 0.028203625231981277

step: 20, loss: 0.0011847828282043338

step: 30, loss: 0.010636034421622753

step: 40, loss: 0.001994851278141141

step: 50, loss: 0.06524886935949326

step: 60, loss: 0.0004916097968816757

step: 70, loss: 0.0017007613787427545

step: 80, loss: 0.0009920833399519324

step: 90, loss: 0.034088943153619766

step: 100, loss: 0.008503897115588188

step: 110, loss: 0.004090732894837856

step: 120, loss: 0.00027571909595280886

step: 130, loss: 0.00923825427889824

step: 140, loss: 0.005794602911919355

step: 150, loss: 0.08688487857580185

step: 160, loss: 0.04171200469136238

step: 170, loss: 0.01058843731880188

step: 180, loss: 0.010342597961425781

step: 190, loss: 0.026668336242437363

step: 200, loss: 0.008357802405953407

step: 210, loss: 0.07980653643608093

step: 220, loss: 0.005970470607280731

step: 230, loss: 0.22685520350933075

step: 240, loss: 0.007239742670208216

step: 250, loss: 0.2430586963891983

step: 260, loss: 0.003947349265217781

step: 270, loss: 0.00822496972978115

step: 280, loss: 0.06469642370939255

step: 290, loss: 0.05334723740816116

step: 300, loss: 0.005584970116615295

step: 310, loss: 0.03837605193257332

step: 320, loss: 0.016373474150896072

step: 330, loss: 0.005537836346775293

step: 340, loss: 0.028610721230506897

step: 350, loss: 0.009495988488197327

step: 360, loss: 0.004364070016890764

step: 370, loss: 0.015741605311632156

step: 380, loss: 0.0005309724365361035

step: 390, loss: 0.0013983292737975717

```

step: 400, loss: 0.025310322642326355
step: 410, loss: 0.02022848092019558
step: 420, loss: 0.004988345317542553
step: 430, loss: 0.0650845617055893
step: 440, loss: 0.012890959158539772
step: 450, loss: 0.008044439367949963
step: 460, loss: 0.00279228831641376
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 470, loss: 0.026406843215227127
step: 480, loss: 0.11307043582201004
step: 490, loss: 0.06795819103717804
step: 500, loss: 0.027499500662088394
step: 510, loss: 0.07388601452112198
step: 520, loss: 0.18033774197101593
step: 530, loss: 0.006432600785046816
/usr/local/lib/python3.7/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=0.9184430027803522, f1=0.9111111111111111,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011259259656071663
step: 10, loss: 0.014430507086217403
step: 20, loss: 0.009627269580960274
step: 30, loss: 0.00036947146872989833
step: 40, loss: 0.0020079908426851034
step: 50, loss: 0.018229100853204727
step: 60, loss: 0.12347866594791412
step: 70, loss: 0.019194088876247406
step: 80, loss: 0.000709387066308409
step: 90, loss: 0.01730933226644993
step: 100, loss: 0.0124126635491848
step: 110, loss: 0.0015604222426190972
step: 120, loss: 0.0007556519703939557
step: 130, loss: 0.0006119221216067672
step: 140, loss: 0.0248943530023098
step: 150, loss: 0.0019758413545787334
step: 160, loss: 0.008183316327631474
step: 170, loss: 0.016103321686387062
step: 180, loss: 0.0011633114190772176
step: 190, loss: 0.012099655345082283
step: 200, loss: 0.0008043827838264406
step: 210, loss: 0.0009214807651005685
step: 220, loss: 0.03143446147441864
step: 230, loss: 0.01515775267034769
step: 240, loss: 0.13566987216472626

```

```

step: 250, loss: 0.0030285411048680544
step: 260, loss: 0.005179464817047119
step: 270, loss: 0.04240249842405319
step: 280, loss: 0.03282366320490837
step: 290, loss: 0.019243117421865463
step: 300, loss: 0.00538916140794754
step: 310, loss: 0.0005949695478193462
step: 320, loss: 0.00967409834265709
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 330, loss: 0.004135778173804283
step: 340, loss: 0.0054864464327692986
step: 350, loss: 0.009542322717607021
step: 360, loss: 0.009919105097651482
step: 370, loss: 0.020972643047571182
step: 380, loss: 0.011542077176272869
step: 390, loss: 0.0002815184707287699
step: 400, loss: 0.0013202282134443521
step: 410, loss: 0.0006294830818660557
step: 420, loss: 0.07810048013925552
step: 430, loss: 0.0009386423625983298
step: 440, loss: 0.0049060313031077385
step: 450, loss: 0.006968605797737837
step: 460, loss: 0.13293731212615967
step: 470, loss: 0.029359495267271996
step: 480, loss: 0.021711237728595734
step: 490, loss: 0.0037637706845998764
step: 500, loss: 0.036442384123802185
step: 510, loss: 0.06823689490556717
step: 520, loss: 0.006009566597640514
step: 530, loss: 0.03209169954061508
/usr/local/lib/python3.7/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=0.9219924812030075, f1=0.9141248240262787,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038415638264268637
step: 10, loss: 0.0004506949335336685
step: 20, loss: 0.13600262999534607
step: 30, loss: 0.0038290026132017374
step: 40, loss: 0.010466529987752438
step: 50, loss: 0.0032597950194031
step: 60, loss: 0.0009400260751135647
step: 70, loss: 0.0006095021381042898
step: 80, loss: 0.0027677295729517937
step: 90, loss: 0.00021338473015930504

```

```
step: 100, loss: 0.0034165570978075266
step: 110, loss: 0.0016427893424406648
step: 120, loss: 0.004648773465305567
step: 130, loss: 0.0014181461883708835
step: 140, loss: 0.01833917945623398
step: 150, loss: 0.0036176894791424274
step: 160, loss: 0.005746215116232634
step: 170, loss: 0.0034601411316543818
step: 180, loss: 0.003710476215928793
step: 190, loss: 0.03165451064705849
step: 200, loss: 0.0011129735503345728
step: 210, loss: 0.041695088148117065
step: 220, loss: 0.0015054827090352774
step: 230, loss: 0.000605646229814738
step: 240, loss: 0.0014108173782005906
step: 250, loss: 0.0004438794276211411
step: 260, loss: 0.00016106519615277648
step: 270, loss: 0.010388505645096302
step: 280, loss: 0.0008250685059465468
step: 290, loss: 0.0019006183138117194
step: 300, loss: 0.0007457399624399841
step: 310, loss: 0.0004645981243811548
step: 320, loss: 0.00014628680946771055
step: 330, loss: 0.00017625506734475493
step: 340, loss: 0.3354082405567169
step: 350, loss: 0.08591793477535248
step: 360, loss: 0.02176196686923504
step: 370, loss: 0.008821824565529823
step: 380, loss: 0.0004673159564845264
step: 390, loss: 0.0030526253394782543
step: 400, loss: 0.0019930130802094936
step: 410, loss: 0.007937571033835411
step: 420, loss: 0.0012885527685284615
step: 430, loss: 0.12044724076986313
step: 440, loss: 0.007809057831764221
step: 450, loss: 0.0015282963868230581
step: 460, loss: 0.0004584007547236979
step: 470, loss: 0.033008839935064316
step: 480, loss: 0.0043427306227386
step: 490, loss: 0.003530388930812478
step: 500, loss: 0.0006111947586759925
step: 510, loss: 0.0019191757310181856
step: 520, loss: 0.006488231010735035
step: 530, loss: 0.005743095185607672
/usr/local/lib/python3.7/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=0.9204119850187266, f1=0.9095149253731343,
```



```
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014867506979499012
step: 10, loss: 0.03390058875083923
step: 20, loss: 0.006252777762711048
step: 30, loss: 0.009573184885084629
step: 40, loss: 0.016746852546930313
step: 50, loss: 0.0015157908201217651
step: 60, loss: 0.0003732129989657551
step: 70, loss: 0.00028929111431352794
step: 80, loss: 0.0024271158035844564
step: 90, loss: 0.00031300634145736694
step: 100, loss: 0.0001271791843464598
step: 110, loss: 0.00020616964320652187
step: 120, loss: 0.003692885860800743
step: 130, loss: 0.0012083754409104586
step: 140, loss: 0.00033761499798856676
step: 150, loss: 0.017765067517757416
step: 160, loss: 0.009878742508590221
step: 170, loss: 0.00106394337490201
step: 180, loss: 0.0003376344102434814
step: 190, loss: 0.0019940368365496397
step: 200, loss: 0.00040119694313034415
step: 210, loss: 0.008541190065443516
step: 220, loss: 0.00022601558885071427
step: 230, loss: 0.0011295683216303587
step: 240, loss: 0.0008802691008895636
step: 250, loss: 0.025172779336571693
step: 260, loss: 0.0009747175499796867
step: 270, loss: 0.00013279034465085715
step: 280, loss: 0.0005889395833946764
step: 290, loss: 0.0001025043020490557
step: 300, loss: 0.25071442127227783
step: 310, loss: 0.00020165419846307486
step: 320, loss: 0.02247069776058197
step: 330, loss: 0.0002877497172448784
step: 340, loss: 0.0003545338404364884
step: 350, loss: 0.004553951323032379
step: 360, loss: 0.0016085507813841105
step: 370, loss: 0.0023332107812166214
step: 380, loss: 0.018730411306023598
step: 390, loss: 0.0014841896481812
step: 400, loss: 0.0006484621553681791
step: 410, loss: 0.00046631123404949903
step: 420, loss: 0.0013130201259627938
step: 430, loss: 0.00032575157820247114
```

```

step: 440, loss: 0.0016825892962515354
step: 450, loss: 0.001059907954186201
step: 460, loss: 0.001159211155027151
step: 470, loss: 0.003919443115592003
step: 480, loss: 0.06684839725494385
step: 490, loss: 0.0015378267271444201
step: 500, loss: 0.021573642268776894
step: 510, loss: 0.001519518904387951
step: 520, loss: 0.010705350898206234
step: 530, loss: 0.0017991483910009265
/usr/local/lib/python3.7/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=0.9246511627906977, f1=0.9090909090909092,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016714130470063537
step: 10, loss: 0.0006173359579406679
step: 20, loss: 0.0006469584186561406
step: 30, loss: 0.00013029453111812472
step: 40, loss: 0.0004040873027406633
step: 50, loss: 0.0015879437560215592
step: 60, loss: 0.00017573987133800983
step: 70, loss: 0.00012610381236299872
step: 80, loss: 0.0009466487099416554
step: 90, loss: 0.0015323205152526498
step: 100, loss: 0.00029330889810808003
step: 110, loss: 0.0004958404460921884
step: 120, loss: 0.0020177369005978107
step: 130, loss: 0.0018410453340038657
step: 140, loss: 0.002190760802477598
step: 150, loss: 0.00042789540020748973
step: 160, loss: 0.14524757862091064
step: 170, loss: 0.003348702099174261
step: 180, loss: 0.010691553354263306
step: 190, loss: 0.005035657435655594
step: 200, loss: 0.002593335695564747
step: 210, loss: 0.005002851597964764
step: 220, loss: 0.003439251333475113
step: 230, loss: 0.0005078867543488741
step: 240, loss: 0.00035655396641232073
step: 250, loss: 0.00020209782815072685
step: 260, loss: 5.099350164528005e-05
step: 270, loss: 0.0010871805716305971
step: 280, loss: 0.12490841001272202
step: 290, loss: 0.0008265124633908272

```

```

step: 300, loss: 0.003196248086169362
step: 310, loss: 0.041697002947330475
step: 320, loss: 0.003212974639609456
step: 330, loss: 0.0003239848301745951
step: 340, loss: 0.0007042799261398613
step: 350, loss: 0.00020958912500645965
step: 360, loss: 0.0002984145248774439
step: 370, loss: 0.00010592663602437824
step: 380, loss: 0.007140308618545532
step: 390, loss: 0.05474354699254036
step: 400, loss: 0.013803962618112564
step: 410, loss: 0.0004100946243852377
step: 420, loss: 0.0004991780151613057
step: 430, loss: 0.00010490623390069231
step: 440, loss: 0.0018744246335700154
step: 450, loss: 8.601044828537852e-05
step: 460, loss: 0.00010475435556145385
step: 470, loss: 0.00024303716782014817
step: 480, loss: 5.097725443192758e-05
step: 490, loss: 0.0004989289445802569
step: 500, loss: 0.00010599527740851045
step: 510, loss: 0.00020238345314282924
step: 520, loss: 0.0003316465299576521
step: 530, loss: 0.00018537462165113539
/usr/local/lib/python3.7/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=0.912739150723285, f1=0.9108635097493036,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00042649658280424774
step: 10, loss: 7.33766210032627e-05
step: 20, loss: 0.0004519375797826797
step: 30, loss: 0.002593549434095621
step: 40, loss: 0.0001229705521836877
step: 50, loss: 8.718219760339707e-05
step: 60, loss: 4.5271634007804096e-05
step: 70, loss: 0.014197969809174538
step: 80, loss: 0.0002422226534690708
step: 90, loss: 7.04526828485541e-05
step: 100, loss: 0.00010178719094255939
step: 110, loss: 0.00015509602962993085
step: 120, loss: 0.0008270168327726424
step: 130, loss: 0.041889533400535583
step: 140, loss: 0.03189496323466301
step: 150, loss: 0.0004198001406621188

```

```

step: 160, loss: 0.001884213648736477
step: 170, loss: 0.009088393300771713
step: 180, loss: 0.000565598311368376
step: 190, loss: 0.00014845664554741234
step: 200, loss: 0.000304271699860692
step: 210, loss: 0.00032206744072027504
step: 220, loss: 0.0005223358166404068
step: 230, loss: 8.815452747512609e-05
step: 240, loss: 0.0028557749465107918
step: 250, loss: 0.0008612078963778913
step: 260, loss: 0.003977011889219284
step: 270, loss: 3.525875217746943e-05
step: 280, loss: 9.610191045794636e-05
step: 290, loss: 9.223592496709898e-05
step: 300, loss: 3.5756831493927166e-05
step: 310, loss: 0.0004322721215430647
step: 320, loss: 0.0003398796543478966
step: 330, loss: 9.862925071502104e-05
step: 340, loss: 3.352611747686751e-05
step: 350, loss: 0.005008072126656771
step: 360, loss: 4.9403886805521324e-05
step: 370, loss: 0.003438948653638363
step: 380, loss: 8.660833555040881e-05
step: 390, loss: 0.0002300425840076059
step: 400, loss: 0.031503550708293915
step: 410, loss: 0.00016708651673980057
step: 420, loss: 2.9514252673834562e-05
step: 430, loss: 4.9157097237184644e-05
step: 440, loss: 0.0006007355987094343
step: 450, loss: 2.868788033083547e-05
step: 460, loss: 3.083308911300264e-05
step: 470, loss: 3.128967364318669e-05
step: 480, loss: 0.0004645706794690341
step: 490, loss: 0.00021663862571585923
step: 500, loss: 0.00022310539497993886
step: 510, loss: 6.935677083674818e-05
step: 520, loss: 6.534886779263616e-05
step: 530, loss: 7.573691254947335e-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=0.9205453690644099, f1=0.9158530915853093,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002067928435280919
step: 10, loss: 0.04984280467033386

```

step: 20, loss: 3.3248212275793776e-05
step: 30, loss: 0.011095286346971989
step: 40, loss: 0.014656639657914639
step: 50, loss: 0.010232078842818737
step: 60, loss: 4.96519060106948e-05
step: 70, loss: 5.754476660513319e-05
step: 80, loss: 0.0010301638394594193
step: 90, loss: 8.399381476920098e-05
step: 100, loss: 3.756740625249222e-05
step: 110, loss: 8.784302190179005e-05
step: 120, loss: 9.745029819896445e-05
step: 130, loss: 3.087371078436263e-05
step: 140, loss: 0.00021034265228081495
step: 150, loss: 2.7201193006476387e-05
step: 160, loss: 5.8042372984346e-05
step: 170, loss: 0.00031023877090774477
step: 180, loss: 0.001350087346509099
step: 190, loss: 0.00018661897047422826
step: 200, loss: 0.0021551884710788727
step: 210, loss: 0.0004460710915736854
step: 220, loss: 0.003167755901813507
step: 230, loss: 0.00020122916612308472
step: 240, loss: 6.155677692731842e-05
step: 250, loss: 7.69945327192545e-05
step: 260, loss: 0.00021212564024608582
step: 270, loss: 9.673328895587474e-05
step: 280, loss: 0.00214194948785007
step: 290, loss: 2.5591671146685258e-05
step: 300, loss: 2.4649449187563732e-05
step: 310, loss: 2.2537553377333097e-05
step: 320, loss: 0.00023045972920954227
step: 330, loss: 8.095172233879566e-05
step: 340, loss: 0.00019244976283516735
step: 350, loss: 0.0014606531476601958
step: 360, loss: 0.00012108897499274462
step: 370, loss: 0.002158563816919923
step: 380, loss: 0.00011504756548674777
step: 390, loss: 0.0027866056188941
step: 400, loss: 5.597628478426486e-05
step: 410, loss: 0.00033000303665176034
step: 420, loss: 0.0012476493138819933
step: 430, loss: 0.00024324825790245086
step: 440, loss: 0.006242386996746063
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 450, loss: 0.0010609725723043084
step: 460, loss: 4.67174977529794e-05
step: 470, loss: 0.002597033977508545
step: 480, loss: 0.00034503822098486125

```

step: 490, loss: 0.0009580937912687659
step: 500, loss: 0.00223310268484056
step: 510, loss: 4.669924601330422e-05
step: 520, loss: 6.40804209979251e-05
step: 530, loss: 0.0026158581022173166
/usr/local/lib/python3.7/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=0.9180327868852459, f1=0.9149232914923291,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018357131630182266
step: 10, loss: 0.0003927650104742497
step: 20, loss: 0.005908928345888853
step: 30, loss: 0.0003872393863275647
step: 40, loss: 0.00010647383896866813
step: 50, loss: 0.0006494352710433304
step: 60, loss: 0.00036462536081671715
step: 70, loss: 0.0005380547372624278
step: 80, loss: 0.00021875265520066023
step: 90, loss: 0.00011829185677925125
step: 100, loss: 0.0001397558080498129
step: 110, loss: 0.00011854553304146975
step: 120, loss: 0.00034669210435822606
step: 130, loss: 0.00013089634012430906
step: 140, loss: 0.00016908383986447006
step: 150, loss: 1.8771475879475474e-05
step: 160, loss: 0.00017446895071770996
step: 170, loss: 0.0005102959694340825
step: 180, loss: 0.00018538373115006834
step: 190, loss: 0.0008197808638215065
step: 200, loss: 0.001298550982028246
step: 210, loss: 2.7849640900967643e-05
step: 220, loss: 0.004357457160949707
step: 230, loss: 0.0006578196189366281
step: 240, loss: 0.00013285773457027972
step: 250, loss: 2.8113303415011615e-05
step: 260, loss: 3.182974978699349e-05
step: 270, loss: 0.00019442036864347756
step: 280, loss: 0.00034900737227872014
step: 290, loss: 8.60176223795861e-05
step: 300, loss: 4.790541424881667e-05
step: 310, loss: 3.603275035857223e-05
step: 320, loss: 2.101023892464582e-05
step: 330, loss: 6.3285646319855e-05
step: 340, loss: 6.103983469074592e-05

```

```

step: 350, loss: 1.7374522940372117e-05
step: 360, loss: 2.116273572028149e-05
step: 370, loss: 0.00016381131717935205
step: 380, loss: 2.1945252228761092e-05
step: 390, loss: 1.6446972949779592e-05
step: 400, loss: 5.2028564823558554e-05
step: 410, loss: 6.946125358808786e-05
step: 420, loss: 3.2110845495481044e-05
step: 430, loss: 2.1859426851733588e-05
step: 440, loss: 5.640901508741081e-05
step: 450, loss: 0.0007005927618592978
step: 460, loss: 0.0001601131516508758
step: 470, loss: 2.4392646082560532e-05
step: 480, loss: 4.320354855735786e-05
step: 490, loss: 3.08097769448068e-05
step: 500, loss: 5.79151492274832e-05
step: 510, loss: 2.2820451704319566e-05
step: 520, loss: 0.008516148664057255
step: 530, loss: 2.1285726688802242e-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=0.9118483412322276, f1=0.9115582592419279,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.545033385104034e-05
step: 10, loss: 2.8690439648926258e-05
step: 20, loss: 2.801276059472002e-05
step: 30, loss: 4.200956391287036e-05
step: 40, loss: 0.0012879689456894994
step: 50, loss: 0.0002518908877391368
step: 60, loss: 0.000456731126178056
step: 70, loss: 2.2150208678795025e-05
step: 80, loss: 1.880868512671441e-05
step: 90, loss: 2.332329677301459e-05
step: 100, loss: 0.00014460274542216212
step: 110, loss: 3.2362451747758314e-05
step: 120, loss: 2.4407176169916056e-05
step: 130, loss: 8.596352563472465e-05
step: 140, loss: 3.6171870306134224e-05
step: 150, loss: 1.383558992529288e-05
step: 160, loss: 0.0001231275382451713
step: 170, loss: 3.977967207902111e-05
step: 180, loss: 0.000998157076537609
step: 190, loss: 0.00012724175758194178
step: 200, loss: 3.07871559925843e-05

```

```

step: 210, loss: 0.0004316982231102884
step: 220, loss: 3.801032289629802e-05
step: 230, loss: 1.8212564100394957e-05
step: 240, loss: 0.0005341109936125576
step: 250, loss: 6.949709495529532e-05
step: 260, loss: 0.00013072283763904124
step: 270, loss: 0.00026644623721949756
step: 280, loss: 0.0001997221406782046
step: 290, loss: 8.08191834948957e-05
step: 300, loss: 0.00018782315601129085
step: 310, loss: 4.472443833947182e-05
step: 320, loss: 2.6242370950058103e-05
step: 330, loss: 4.1104183765128255e-05
step: 340, loss: 5.48840471310541e-05
step: 350, loss: 2.8515974918263964e-05
step: 360, loss: 5.538127152249217e-05
step: 370, loss: 1.9106513718725182e-05
step: 380, loss: 3.700470915646292e-05
step: 390, loss: 2.0060246242792346e-05
step: 400, loss: 0.00021557418222073466
step: 410, loss: 0.000714583380613476
step: 420, loss: 0.00010434497380629182
step: 430, loss: 0.00016715130186639726
step: 440, loss: 1.8078497305396013e-05
step: 450, loss: 6.633823795709759e-05
step: 460, loss: 1.4636461855843663e-05
step: 470, loss: 2.5621029635658488e-05
step: 480, loss: 5.3854364523431286e-05
step: 490, loss: 4.1352850530529395e-05
step: 500, loss: 1.5090975466591772e-05
step: 510, loss: 0.000527373340446502
step: 520, loss: 1.9303955923533067e-05
step: 530, loss: 2.469740684318822e-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=0.9167855444602948, f1=0.9092618711800659,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0022030407562851906
step: 10, loss: 1.511703976575518e-05
step: 20, loss: 1.09783477455494e-05
step: 30, loss: 3.669544094009325e-05
step: 40, loss: 1.3030941772740334e-05
step: 50, loss: 2.9289298254298046e-05
step: 60, loss: 0.000136022106744349

```


step: 70, loss: 2.1579739041044377e-05
step: 80, loss: 0.00010182994446950033
step: 90, loss: 1.990704913623631e-05
step: 100, loss: 2.479413706168998e-05
step: 110, loss: 1.2922911082569044e-05
step: 120, loss: 0.00024913097149692476
step: 130, loss: 4.441468990989961e-05
step: 140, loss: 0.00016415906429756433
step: 150, loss: 1.8849503248929977e-05
step: 160, loss: 1.4368285519594792e-05
step: 170, loss: 0.0006751790642738342
step: 180, loss: 7.090876169968396e-05
step: 190, loss: 4.797809378942475e-05
step: 200, loss: 0.0002177448768634349
step: 210, loss: 5.424793198471889e-05
step: 220, loss: 1.6625695934635587e-05
step: 230, loss: 1.702424015093129e-05
step: 240, loss: 0.005583968944847584
step: 250, loss: 3.609544364735484e-05
step: 260, loss: 1.2028862329316325e-05
step: 270, loss: 0.00017362984362989664
step: 280, loss: 2.9950295356684364e-05
step: 290, loss: 2.1147236111573875e-05
step: 300, loss: 1.887191683636047e-05
step: 310, loss: 7.13862682459876e-05
step: 320, loss: 0.000194824009668082
step: 330, loss: 4.6030330850044265e-05
step: 340, loss: 1.6226989828282967e-05
step: 350, loss: 0.008098144084215164
step: 360, loss: 1.5884388631093316e-05
step: 370, loss: 2.7401361876400188e-05
step: 380, loss: 1.497178618592443e-05
step: 390, loss: 0.0006920714513398707
step: 400, loss: 0.013057461008429527
step: 410, loss: 6.0428486904129386e-05
step: 420, loss: 1.5504323528148234e-05
step: 430, loss: 9.669783321442083e-05
step: 440, loss: 0.0006827648030593991
step: 450, loss: 0.00029797927709296346
step: 460, loss: 8.320822234963998e-05
step: 470, loss: 3.242158345528878e-05
step: 480, loss: 4.39739778812509e-05
step: 490, loss: 0.0009211646392941475
step: 500, loss: 4.1045139369089156e-05
step: 510, loss: 0.1289994865655899
step: 520, loss: 3.210422437405214e-05
step: 530, loss: 4.279016866348684e-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=0.9050170482221139, f1=0.8794117647058824,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 1.768699257809203e-05
step: 10, loss: 9.24611413211096e-06
step: 20, loss: 7.038878538878635e-05
step: 30, loss: 3.497257421258837e-05
step: 40, loss: 1.896874346130062e-05
step: 50, loss: 0.030894627794623375
step: 60, loss: 6.0144469898659736e-05
step: 70, loss: 9.506958303973079e-05
step: 80, loss: 0.0001722960005281493
step: 90, loss: 9.551404946250841e-05
step: 100, loss: 0.009285589680075645
step: 110, loss: 4.2557090637274086e-05
step: 120, loss: 2.8157475753687322e-05
step: 130, loss: 3.1239505915436894e-05
step: 140, loss: 7.571840978926048e-05
step: 150, loss: 1.4867076970404014e-05
step: 160, loss: 4.0718285163166e-05
step: 170, loss: 0.0008979073609225452
step: 180, loss: 4.957486453349702e-05
step: 190, loss: 1.8372287740930915e-05
step: 200, loss: 2.511756247258745e-05
step: 210, loss: 1.4137289326754399e-05
step: 220, loss: 4.5621698518516496e-05
step: 230, loss: 6.634694000240415e-05
step: 240, loss: 1.3351207599043846e-05
step: 250, loss: 2.1855710656382143e-05
step: 260, loss: 0.00019112277368549258
step: 270, loss: 0.00023554319341201335
step: 280, loss: 1.091501417249674e-05
step: 290, loss: 3.907713835360482e-05
step: 300, loss: 4.811911639990285e-05
step: 310, loss: 7.286226173164323e-05
step: 320, loss: 2.7356993086868897e-05
step: 330, loss: 3.6551344237523153e-05
step: 340, loss: 0.00015740326489321887
step: 350, loss: 2.534527811803855e-05
step: 360, loss: 4.094117321074009e-05
step: 370, loss: 1.0110375114891212e-05
step: 380, loss: 6.502964242827147e-05
step: 390, loss: 0.0003956523723900318
step: 400, loss: 3.240141086280346e-05

```

```

step: 410, loss: 7.005450606811792e-05
step: 420, loss: 1.4450117305386811e-05
step: 430, loss: 1.448373313905904e-05
step: 440, loss: 3.084411218878813e-05
step: 450, loss: 5.0080696382792667e-05
step: 460, loss: 1.8145412468584254e-05
step: 470, loss: 1.751956733642146e-05
step: 480, loss: 1.2855805834988132e-05
step: 490, loss: 3.344158176332712e-05
step: 500, loss: 1.2501946912379935e-05
step: 510, loss: 0.00012879702262580395
step: 520, loss: 0.00023554485233034939
step: 530, loss: 0.000615019875112921
/usr/local/lib/python3.7/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=0.9203624225083452, f1=0.9053732762719924,
best_f1=0.9136490250696379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011675584391923621
step: 10, loss: 6.656504410784692e-05
step: 20, loss: 1.2516798960859887e-05
step: 30, loss: 0.0019949893467128277
step: 40, loss: 0.06550329923629761
step: 50, loss: 6.963990745134652e-05
step: 60, loss: 1.036740741255926e-05
step: 70, loss: 0.0005621804157271981
step: 80, loss: 6.962265615584329e-05
step: 90, loss: 1.518408225820167e-05
step: 100, loss: 1.1432798601163086e-05
step: 110, loss: 1.9243736460339278e-05
step: 120, loss: 3.0077004339545965e-05
step: 130, loss: 0.005312950350344181
step: 140, loss: 2.787277480820194e-05
step: 150, loss: 1.5306995919672772e-05
step: 160, loss: 1.3377278264670167e-05
step: 170, loss: 0.00490127969533205
step: 180, loss: 1.4025443306309171e-05
step: 190, loss: 1.3261887033877429e-05
step: 200, loss: 1.6688562027411535e-05
step: 210, loss: 6.225573451956734e-05
step: 220, loss: 1.3708761798625346e-05
step: 230, loss: 2.414016125840135e-05
step: 240, loss: 0.0002544661983847618
step: 250, loss: 2.2141435692901723e-05
step: 260, loss: 1.4457610632234719e-05

```

```

step: 270, loss: 1.204747604788281e-05
step: 280, loss: 1.301975680689793e-05
step: 290, loss: 1.0576027307251934e-05
step: 300, loss: 1.1440251910244115e-05
step: 310, loss: 2.5888723030220717e-05
step: 320, loss: 1.1097522474301513e-05
step: 330, loss: 4.899277337244712e-05
step: 340, loss: 1.1276320037723053e-05
step: 350, loss: 2.2834061383036897e-05
step: 360, loss: 3.8504636904690415e-05
step: 370, loss: 1.043817974277772e-05
step: 380, loss: 0.00013605623098555952
step: 390, loss: 0.0001284056343138218
step: 400, loss: 2.4805040084174834e-05
step: 410, loss: 2.1672547518392093e-05
step: 420, loss: 2.2961377908359282e-05
step: 430, loss: 2.357243101869244e-05
step: 440, loss: 0.0002953716611955315
step: 450, loss: 9.559031241224147e-06
step: 460, loss: 0.00022805640765000135
step: 470, loss: 0.004405771382153034
step: 480, loss: 9.044957550941035e-06
step: 490, loss: 1.9973875168943778e-05
step: 500, loss: 1.2937761312059592e-05
step: 510, loss: 1.64578013936989e-05
step: 520, loss: 2.2536398319061846e-05
step: 530, loss: 2.6427716875332408e-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=0.9179559306141585, f1=0.911970190964136,
best_f1=0.9136490250696379
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 [00:29, 192.79it/s]
load_f1 = 0.9218106995884773
real_f1 = 0.9189435336976319
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, 190.95it/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: 0.5611923336982727
```

```
step: 10, loss: 0.38644158840179443
```

```

step: 20, loss: 0.37713268399238586
step: 30, loss: 0.3179284632205963
step: 40, loss: 0.17053601145744324
step: 50, loss: 0.44946733117103577
step: 60, loss: 0.3093242049217224
step: 70, loss: 0.2078080177307129
step: 80, loss: 0.20134606957435608
step: 90, loss: 0.4359448552131653
step: 100, loss: 0.49429771304130554
step: 110, loss: 0.2555061876773834
step: 120, loss: 0.22763706743717194
step: 130, loss: 0.27882811427116394
step: 140, loss: 0.27412712574005127
step: 150, loss: 0.3425874412059784
step: 160, loss: 0.29824167490005493
step: 170, loss: 0.2915915548801422
step: 180, loss: 0.11215002834796906
step: 190, loss: 0.22859713435173035
step: 200, loss: 0.27091020345687866
step: 210, loss: 0.26401573419570923
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5330490405117271, f1=0.5341880341880342,
best_f1=0.5341880341880342
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16162054240703583
step: 10, loss: 0.24444811046123505
step: 20, loss: 0.22366714477539062
step: 30, loss: 0.319555401802063
step: 40, loss: 0.15789036452770233
step: 50, loss: 0.16771237552165985
step: 60, loss: 0.3360796868801117
step: 70, loss: 0.15536245703697205
step: 80, loss: 0.12254005670547485
step: 90, loss: 0.14541083574295044
step: 100, loss: 0.03212235867977142
step: 110, loss: 0.11083339154720306
step: 120, loss: 0.12345631420612335
step: 130, loss: 0.04446383938193321
step: 140, loss: 0.23097017407417297
step: 150, loss: 0.2991699278354645

```

```

step: 160, loss: 0.2259792983531952
step: 170, loss: 0.11210542917251587
step: 180, loss: 0.21717596054077148
step: 190, loss: 0.20722271502017975
step: 200, loss: 0.0963422879576683
step: 210, loss: 0.1976414918899536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5647969052224371, f1=0.540650406504065,
best_f1=0.540650406504065
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0786239355802536
step: 10, loss: 0.25027549266815186
step: 20, loss: 0.1248626783490181
step: 30, loss: 0.21035796403884888
step: 40, loss: 0.0760589987039566
step: 50, loss: 0.12855404615402222
step: 60, loss: 0.22699202597141266
step: 70, loss: 0.1573581099510193
step: 80, loss: 0.17279882729053497
step: 90, loss: 0.060598116368055344
step: 100, loss: 0.23000800609588623
step: 110, loss: 0.11647797375917435
step: 120, loss: 0.14938409626483917
step: 130, loss: 0.12324649840593338
step: 140, loss: 0.14276239275932312
step: 150, loss: 0.25406700372695923
step: 160, loss: 0.033351533114910126
step: 170, loss: 0.11895205825567245
step: 180, loss: 0.13711927831172943
step: 190, loss: 0.24073848128318787
step: 200, loss: 0.050112321972846985
step: 210, loss: 0.13689330220222473
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5778688524590164, f1=0.5656565656565656,

```



```

best_f1=0.5656565656565656
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3278016746044159
step: 10, loss: 0.16932815313339233
step: 20, loss: 0.10403262823820114
step: 30, loss: 0.26789772510528564
step: 40, loss: 0.044133733958005905
step: 50, loss: 0.17485059797763824
step: 60, loss: 0.23218899965286255
step: 70, loss: 0.22304223477840424
step: 80, loss: 0.1948365420103073
step: 90, loss: 0.07992059737443924
step: 100, loss: 0.40223580598831177
step: 110, loss: 0.1328185498714447
step: 120, loss: 0.04681597277522087
step: 130, loss: 0.19121988117694855
step: 140, loss: 0.21643094718456268
step: 150, loss: 0.05978627875447273
step: 160, loss: 0.03636246547102928
step: 170, loss: 0.06734732538461685
step: 180, loss: 0.3515089154243469
step: 190, loss: 0.10944835096597672
step: 200, loss: 0.38125428557395935
step: 210, loss: 0.0996648520231247
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6061776061776061, f1=0.5568627450980392,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16391387581825256
step: 10, loss: 0.034398213028907776
step: 20, loss: 0.34440550208091736
step: 30, loss: 0.06831744313240051
step: 40, loss: 0.1328437179327011
step: 50, loss: 0.13717681169509888
step: 60, loss: 0.18958395719528198
step: 70, loss: 0.24803631007671356
step: 80, loss: 0.07886286079883575
step: 90, loss: 0.1498102992773056

```

```

step: 100, loss: 0.04501115158200264
step: 110, loss: 0.19307257235050201
step: 120, loss: 0.2090694010257721
step: 130, loss: 0.11748191714286804
step: 140, loss: 0.09425688534975052
step: 150, loss: 0.05148768052458763
step: 160, loss: 0.12288381159305573
step: 170, loss: 0.07285480946302414
step: 180, loss: 0.06214740872383118
step: 190, loss: 0.08364133536815643
step: 200, loss: 0.05173291265964508
step: 210, loss: 0.11025509238243103
/usr/local/lib/python3.7/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=0.5708812260536399, f1=0.5557729941291586,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.046704329550266266
step: 10, loss: 0.1789127141237259
step: 20, loss: 0.05912996456027031
step: 30, loss: 0.011254316195845604
step: 40, loss: 0.09306842088699341
step: 50, loss: 0.01864558644592762
step: 60, loss: 0.18297481536865234
step: 70, loss: 0.011721691116690636
step: 80, loss: 0.04639827460050583
step: 90, loss: 0.23857198655605316
step: 100, loss: 0.005800997838377953
step: 110, loss: 0.0953008309006691
step: 120, loss: 0.035374678671360016
step: 130, loss: 0.0927596166729927
step: 140, loss: 0.0954502522945404
step: 150, loss: 0.026092084124684334
step: 160, loss: 0.02586241438984871
step: 170, loss: 0.1270386278629303
step: 180, loss: 0.12866972386837006
step: 190, loss: 0.07058203220367432
step: 200, loss: 0.01598399318754673
step: 210, loss: 0.06956134736537933
/usr/local/lib/python3.7/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=0.556420233463035, f1=0.569672131147541,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0409090481698513
step: 10, loss: 0.10740915685892105
step: 20, loss: 0.009524459950625896
step: 30, loss: 0.0331883504986763
step: 40, loss: 0.08799392729997635
step: 50, loss: 0.10713519901037216
step: 60, loss: 0.027621876448392868
step: 70, loss: 0.04417441412806511
step: 80, loss: 0.034628208726644516
step: 90, loss: 0.16476179659366608
step: 100, loss: 0.00431300513446331
step: 110, loss: 0.036298610270023346
step: 120, loss: 0.10140202194452286
step: 130, loss: 0.14839857816696167
step: 140, loss: 0.01828221045434475
step: 150, loss: 0.06589596718549728
step: 160, loss: 0.029395654797554016
step: 170, loss: 0.02244126982986927
step: 180, loss: 0.142632856965065
step: 190, loss: 0.0638015866279602
step: 200, loss: 0.032267976552248
step: 210, loss: 0.10985866189002991
/usr/local/lib/python3.7/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=0.5291005291005292, f1=0.5237258347978909,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012907147407531738
step: 10, loss: 0.01829376071691513
step: 20, loss: 0.01357145607471466
step: 30, loss: 0.21957461535930634
step: 40, loss: 0.02363421395421028
step: 50, loss: 0.024418950080871582
step: 60, loss: 0.009044647216796875
step: 70, loss: 0.19680695235729218
step: 80, loss: 0.06637648493051529
step: 90, loss: 0.1169770136475563
step: 100, loss: 0.028216486796736717
step: 110, loss: 0.05877493694424629
step: 120, loss: 0.001754534081555903
step: 130, loss: 0.03900238871574402
step: 140, loss: 0.0484958253800869
step: 150, loss: 0.07927308976650238

```

```

step: 160, loss: 0.03095635212957859
step: 170, loss: 0.01599019020795822
step: 180, loss: 0.08596393465995789
step: 190, loss: 0.09014039486646652
step: 200, loss: 0.002551599871367216
step: 210, loss: 0.10988692194223404
/usr/local/lib/python3.7/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=0.5546558704453441, f1=0.5543710021321961,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.042269814759492874
step: 10, loss: 0.05633142217993736
step: 20, loss: 0.0028205791022628546
step: 30, loss: 0.021612150594592094
step: 40, loss: 0.054686516523361206
step: 50, loss: 0.06672314554452896
step: 60, loss: 0.10029994696378708
step: 70, loss: 0.06999605894088745
step: 80, loss: 0.039950739592313766
step: 90, loss: 0.0023049444425851107
step: 100, loss: 0.03955827280879021
step: 110, loss: 0.020106468349695206
step: 120, loss: 0.0011533934157341719
step: 130, loss: 0.08392435312271118
step: 140, loss: 0.15580138564109802
step: 150, loss: 0.19150902330875397
step: 160, loss: 0.010594744235277176
step: 170, loss: 0.1688295602798462
step: 180, loss: 0.09139091521501541
step: 190, loss: 0.010592738166451454
step: 200, loss: 0.022085247561335564
step: 210, loss: 0.007988912984728813
/usr/local/lib/python3.7/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=0.5361552028218695, f1=0.5261261261261261,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008069775998592377
step: 10, loss: 0.1595674455165863
step: 20, loss: 0.005499511957168579
step: 30, loss: 0.09049864113330841

```

```

step: 40, loss: 0.004341701976954937
step: 50, loss: 0.010579517111182213
step: 60, loss: 0.04607633873820305
step: 70, loss: 0.04097166657447815
step: 80, loss: 0.06995012611150742
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 90, loss: 0.03178061917424202
step: 100, loss: 0.12314648926258087
step: 110, loss: 0.0015348304295912385
step: 120, loss: 0.0051179141737520695
step: 130, loss: 0.21108034253120422
step: 140, loss: 0.0036767160054296255
step: 150, loss: 0.05806631222367287
step: 160, loss: 0.047275856137275696
step: 170, loss: 0.005568955559283495
step: 180, loss: 0.1410343199968338
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.18182332813739777
step: 200, loss: 0.06139504164457321
step: 210, loss: 0.1595248132944107
/usr/local/lib/python3.7/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=0.5150375939849624, f1=0.5454545454545454,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0453801229596138
step: 10, loss: 0.06895886361598969
step: 20, loss: 0.06751503795385361
step: 30, loss: 0.0040206327103078365
step: 40, loss: 0.0077078803442418575
step: 50, loss: 0.0010736518306657672
step: 60, loss: 0.1951902210712433
step: 70, loss: 0.006906111259013414
step: 80, loss: 0.3245992362499237
step: 90, loss: 0.012729942798614502
step: 100, loss: 0.2924582064151764
step: 110, loss: 0.012566099874675274
step: 120, loss: 0.15929563343524933
step: 130, loss: 0.016764821484684944
step: 140, loss: 0.0016344841569662094
step: 150, loss: 0.0018078511347994208
step: 160, loss: 0.011700207367539406
step: 170, loss: 0.0247363168746233
step: 180, loss: 0.0037898330483585596
step: 190, loss: 0.01339690014719963

```

```

step: 200, loss: 0.015279043465852737
step: 210, loss: 0.011564263142645359
/usr/local/lib/python3.7/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=0.5530303030303031, f1=0.5490196078431372,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003863443387672305
step: 10, loss: 0.04062485694885254
step: 20, loss: 0.02173774503171444
step: 30, loss: 0.02984962984919548
step: 40, loss: 0.010440164245665073
step: 50, loss: 0.0007443698705174029
step: 60, loss: 0.044044654816389084
step: 70, loss: 0.046471595764160156
step: 80, loss: 0.1461157351732254
step: 90, loss: 0.003963339142501354
step: 100, loss: 0.0011230690870434046
step: 110, loss: 0.023044582456350327
step: 120, loss: 0.09611707180738449
step: 130, loss: 0.011642190627753735
step: 140, loss: 0.001475564669817686
step: 150, loss: 0.004164684098213911
step: 160, loss: 0.030748741701245308
step: 170, loss: 0.0014086613664403558
step: 180, loss: 0.09460410475730896
step: 190, loss: 0.00832638144493103
step: 200, loss: 0.004148244857788086
step: 210, loss: 0.01752704195678234
/usr/local/lib/python3.7/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=0.517453798767967, f1=0.5299145299145299,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005693774204701185
step: 10, loss: 0.0005724240909330547
step: 20, loss: 0.1948850005865097
step: 30, loss: 0.0848473459482193
step: 40, loss: 0.026782728731632233
step: 50, loss: 0.17179670929908752
step: 60, loss: 0.005470363423228264
step: 70, loss: 0.010140486061573029

```

```

step: 80, loss: 0.028449513018131256
step: 90, loss: 0.0018416172824800014
step: 100, loss: 0.003859027987346053
step: 110, loss: 0.017686694860458374
step: 120, loss: 0.10191787779331207
step: 130, loss: 0.0012477375566959381
step: 140, loss: 0.004312962293624878
step: 150, loss: 0.014039658941328526
step: 160, loss: 0.0010092156007885933
step: 170, loss: 0.0017462980467826128
step: 180, loss: 0.04151666536927223
step: 190, loss: 0.0007352022221311927
step: 200, loss: 0.01745576225221157
step: 210, loss: 0.0253006462007761
/usr/local/lib/python3.7/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=0.5233644859813084, f1=0.5330739299610895,
best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016086819814518094
step: 10, loss: 0.0441623218357563
step: 20, loss: 0.1975434273481369
step: 30, loss: 0.03178960457444191
step: 40, loss: 0.0011209631338715553
step: 50, loss: 0.004676549695432186
step: 60, loss: 0.006649536546319723
step: 70, loss: 0.00691120745614171
step: 80, loss: 0.002532718935981393
step: 90, loss: 0.1208195760846138
step: 100, loss: 0.009966066107153893
step: 110, loss: 0.017402222380042076
step: 120, loss: 0.02383848838508129
step: 130, loss: 0.003400550689548254
step: 140, loss: 0.04728928208351135
step: 150, loss: 0.0031866838689893484
step: 160, loss: 0.008638273924589157
step: 170, loss: 0.000620245176833123
step: 180, loss: 0.018532585352659225
step: 190, loss: 0.00390249234624207
step: 200, loss: 0.0112845990806818
step: 210, loss: 0.0021099913865327835
/usr/local/lib/python3.7/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=0.5325443786982248, f1=0.5405405405405405,

```

```

best_f1=0.5568627450980392
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024794195778667927
step: 10, loss: 0.0013929991982877254
step: 20, loss: 0.0047808559611439705
step: 30, loss: 0.008889980614185333
step: 40, loss: 0.0006930950912646949
step: 50, loss: 0.00451296241953969
step: 60, loss: 0.03798713907599449
step: 70, loss: 0.001218621851876378
step: 80, loss: 0.001083742012269795
step: 90, loss: 0.0026793323922902346
step: 100, loss: 0.0018487529596313834
step: 110, loss: 0.04213309660553932
step: 120, loss: 0.003612432163208723
step: 130, loss: 0.0023345653899013996
step: 140, loss: 0.0005168608040548861
step: 150, loss: 0.0007911178399808705
step: 160, loss: 0.012008327059447765
step: 170, loss: 0.011058385483920574
step: 180, loss: 0.0008063867571763694
step: 190, loss: 0.0007665580487810075
step: 200, loss: 0.003915850538760424
step: 210, loss: 0.01838054321706295
/usr/local/lib/python3.7/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=0.5288640595903166, f1=0.535645472061657,
best_f1=0.5568627450980392
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, 450.24it/s]
load_f1 = 0.6020202020202021
real_f1 = 0.6053169734151329
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, 274.64it/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        : 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: 0.5789780616760254

step: 10, loss: 0.36953702569007874

step: 20, loss: 0.28820955753326416

```

step: 30, loss: 0.41517430543899536
step: 40, loss: 0.42503538727760315
step: 50, loss: 0.2819135785102844
step: 60, loss: 0.286545068025589
step: 70, loss: 0.263368159532547
step: 80, loss: 0.24559541046619415
step: 90, loss: 0.2796720564365387
step: 100, loss: 0.3162561357021332
step: 110, loss: 0.5183542966842651
step: 120, loss: 0.20138925313949585
step: 130, loss: 0.14013135433197021
step: 140, loss: 0.13970859348773956
step: 150, loss: 0.1947200745344162
step: 160, loss: 0.11561141908168793
step: 170, loss: 0.2195371389389038
step: 180, loss: 0.04553963989019394
step: 190, loss: 0.1884872317314148
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.610079575596817, f1=0.6351706036745407,
best_f1=0.6351706036745407
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.22974978387355804
step: 10, loss: 0.0707065537571907
step: 20, loss: 0.25796306133270264
step: 30, loss: 0.1384650021791458
step: 40, loss: 0.19323372840881348
step: 50, loss: 0.09259499609470367
step: 60, loss: 0.4000706970691681
step: 70, loss: 0.19665755331516266
step: 80, loss: 0.1810174137353897
step: 90, loss: 0.243983656167984
step: 100, loss: 0.03509033843874931
step: 110, loss: 0.10647857189178467
step: 120, loss: 0.22845697402954102
step: 130, loss: 0.10397869348526001
step: 140, loss: 0.04133663699030876
step: 150, loss: 0.12040731310844421
step: 160, loss: 0.05393365025520325
step: 170, loss: 0.19596907496452332
step: 180, loss: 0.14984510838985443

```

```

step: 190, loss: 0.04124350845813751
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.721763085399449, f1=0.7675675675675676,
best_f1=0.7675675675675676
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0605936236679554
step: 10, loss: 0.2880299389362335
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.13364887237548828
step: 30, loss: 0.08962225168943405
step: 40, loss: 0.057532429695129395
step: 50, loss: 0.15332122147083282
step: 60, loss: 0.03083973191678524
step: 70, loss: 0.10247563570737839
step: 80, loss: 0.0625186339020729
step: 90, loss: 0.03790021315217018
step: 100, loss: 0.15905842185020447
step: 110, loss: 0.10412997007369995
step: 120, loss: 0.1852577030658722
step: 130, loss: 0.01340932585299015
step: 140, loss: 0.04333516210317612
step: 150, loss: 0.13104046881198883
step: 160, loss: 0.043685756623744965
step: 170, loss: 0.12182677537202835
step: 180, loss: 0.16411714255809784
step: 190, loss: 0.1302795708179474
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7403314917127072, f1=0.7639257294429709,
best_f1=0.7639257294429709
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02032320760190487
step: 10, loss: 0.12911422550678253

```

```

step: 20, loss: 0.04686703905463219
step: 30, loss: 0.009608601219952106
step: 40, loss: 0.027042513713240623
step: 50, loss: 0.044224295765161514
step: 60, loss: 0.04998878389596939
step: 70, loss: 0.11743656545877457
step: 80, loss: 0.25848880410194397
step: 90, loss: 0.043066903948783875
step: 100, loss: 0.09558844566345215
step: 110, loss: 0.0035479555372148752
step: 120, loss: 0.059157490730285645
step: 130, loss: 0.228201225399971
step: 140, loss: 0.008473467081785202
step: 150, loss: 0.10354060679674149
step: 160, loss: 0.03795556724071503
step: 170, loss: 0.1544751226902008
step: 180, loss: 0.030785009264945984
step: 190, loss: 0.2095428705215454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7507002801120448, f1=0.7534626038781163,
best_f1=0.7534626038781163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09454938769340515
step: 10, loss: 0.01559077762067318
step: 20, loss: 0.19055798649787903
step: 30, loss: 0.0050519309006631374
step: 40, loss: 0.08499568700790405
step: 50, loss: 0.1076471135020256
step: 60, loss: 0.11141052842140198
step: 70, loss: 0.09650737792253494
step: 80, loss: 0.045030198991298676
step: 90, loss: 0.04471242427825928
step: 100, loss: 0.004418585449457169
step: 110, loss: 0.004532123450189829
step: 120, loss: 0.008919503539800644
step: 130, loss: 0.09201938658952713
step: 140, loss: 0.0058862315490841866
step: 150, loss: 0.032572951167821884
step: 160, loss: 0.01960659958422184
step: 170, loss: 0.0027388008311390877

```

```

step: 180, loss: 0.09577653557062149
step: 190, loss: 0.2008853703737259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7507163323782234, f1=0.7885714285714286,
best_f1=0.7885714285714286
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009264182299375534
step: 10, loss: 0.04268626496195793
step: 20, loss: 0.07872835546731949
step: 30, loss: 0.00243653217330575
step: 40, loss: 0.0165623277425766
step: 50, loss: 0.3007354438304901
step: 60, loss: 0.0690053254365921
step: 70, loss: 0.011880913749337196
step: 80, loss: 0.020938649773597717
step: 90, loss: 0.006715728435665369
step: 100, loss: 0.0008377659250982106
step: 110, loss: 0.08632192015647888
step: 120, loss: 0.10420991480350494
step: 130, loss: 0.0207042433321476
step: 140, loss: 0.1027919352054596
step: 150, loss: 0.022086096927523613
step: 160, loss: 0.149509996175766
step: 170, loss: 0.1432170271873474
step: 180, loss: 0.07654202729463577
step: 190, loss: 0.02561904862523079
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.769230769230769, f1=0.7862796833773087,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004230233374983072
step: 10, loss: 0.013088857755064964

```

```

step: 20, loss: 0.10547413676977158
step: 30, loss: 0.03436967357993126
step: 40, loss: 0.12162870913743973
step: 50, loss: 0.014881747774779797
step: 60, loss: 0.015987563878297806
step: 70, loss: 0.01066247746348381
step: 80, loss: 0.03012465313076973
step: 90, loss: 0.026821885257959366
step: 100, loss: 0.0030695144087076187
step: 110, loss: 0.02125442959368229
step: 120, loss: 0.006660813000053167
step: 130, loss: 0.051666468381881714
step: 140, loss: 0.0053292918018996716
step: 150, loss: 0.19344674050807953
step: 160, loss: 0.09903251379728317
step: 170, loss: 0.08410864323377609
step: 180, loss: 0.0886203944683075
step: 190, loss: 0.0043763150461018085
/usr/local/lib/python3.7/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=0.736, f1=0.7493403693931397, best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03033001720905304
step: 10, loss: 0.13310778141021729
step: 20, loss: 0.002234599320217967
step: 30, loss: 0.0053597684018313885
step: 40, loss: 0.0024021512363106012
step: 50, loss: 0.018101884052157402
step: 60, loss: 0.0007721498841419816
step: 70, loss: 0.015487303026020527
step: 80, loss: 0.0005952980718575418
step: 90, loss: 0.009936174377799034
step: 100, loss: 0.005480511579662561
step: 110, loss: 0.0011917094234377146
step: 120, loss: 0.028845101594924927
step: 130, loss: 0.04500829055905342
step: 140, loss: 0.008952879346907139
step: 150, loss: 0.02405468188226223
step: 160, loss: 0.0015420994022861123
step: 170, loss: 0.014949850738048553
step: 180, loss: 0.10136955976486206
step: 190, loss: 0.03197486326098442
/usr/local/lib/python3.7/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=0.7624309392265193, f1=0.7870619946091645,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026424131356179714
step: 10, loss: 0.004862011875957251
step: 20, loss: 0.003341736737638712
step: 30, loss: 0.05775933340191841
step: 40, loss: 0.002908717840909958
step: 50, loss: 0.00037783378502354026
step: 60, loss: 0.0677393227815628
step: 70, loss: 0.0025871770922094584
step: 80, loss: 0.000566869683098048
step: 90, loss: 0.04441853612661362
step: 100, loss: 0.01001770794391632
step: 110, loss: 0.38785648345947266
step: 120, loss: 0.006342542823404074
step: 130, loss: 0.005050832871347666
step: 140, loss: 0.03435634821653366
step: 150, loss: 0.0013472180580720305
step: 160, loss: 0.0007398475427180529
step: 170, loss: 0.008516873233020306
step: 180, loss: 0.002999192336574197
step: 190, loss: 0.0007945880061015487
/usr/local/lib/python3.7/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=0.7473684210526317, f1=0.7554347826086956,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06827664375305176
step: 10, loss: 0.012598179280757904
step: 20, loss: 0.0011437606299296021
step: 30, loss: 0.00026153732324019074
step: 40, loss: 0.015516798011958599
step: 50, loss: 0.06546138226985931
step: 60, loss: 0.0021441469434648752
step: 70, loss: 0.010609024204313755
step: 80, loss: 0.0027603176422417164
step: 90, loss: 0.006664047483354807
step: 100, loss: 0.004021292086690664
step: 110, loss: 0.003580769058316946
step: 120, loss: 0.05678167939186096
step: 130, loss: 0.000510143639985472
step: 140, loss: 0.019508328288793564

```



```

step: 150, loss: 0.12157556414604187
step: 160, loss: 0.1409148871898651
step: 170, loss: 0.003478032536804676
step: 180, loss: 0.0028793474193662405
step: 190, loss: 0.000399329059291631
/usr/local/lib/python3.7/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=0.7626666666666667, f1=0.7626666666666667,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007972357561811805
step: 10, loss: 0.0011769896373152733
step: 20, loss: 0.02034098654985428
step: 30, loss: 0.00364372949115932
step: 40, loss: 0.003186040325090289
step: 50, loss: 0.009168673306703568
step: 60, loss: 0.004583810921758413
step: 70, loss: 0.004777212161570787
step: 80, loss: 0.0021238457411527634
step: 90, loss: 0.014247671701014042
step: 100, loss: 0.0009058595751412213
step: 110, loss: 0.00044093371252529323
step: 120, loss: 0.0016728575574234128
step: 130, loss: 0.0025869428645819426
step: 140, loss: 0.00088430882897228
step: 150, loss: 0.0002982032601721585
step: 160, loss: 0.0013027189997956157
step: 170, loss: 0.001106219831854105
step: 180, loss: 0.0038884689565747976
step: 190, loss: 0.0013050114503130317
/usr/local/lib/python3.7/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=0.7627118644067796, f1=0.768361581920904,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00039595930138602853
step: 10, loss: 0.0012279037619009614
step: 20, loss: 0.0006379805272445083
step: 30, loss: 0.00043412367813289165
step: 40, loss: 0.000470335507998243
step: 50, loss: 0.019260572269558907
step: 60, loss: 0.004150095861405134

```

```

step: 70, loss: 0.0005889189778827131
step: 80, loss: 0.0008083109278231859
step: 90, loss: 0.001044582691974938
step: 100, loss: 0.0004043552035000175
step: 110, loss: 0.0003799877013079822
step: 120, loss: 0.0031926052179187536
step: 130, loss: 0.005202308762818575
step: 140, loss: 0.0007691237842664123
step: 150, loss: 0.008272187784314156
step: 160, loss: 0.0002163319877581671
step: 170, loss: 0.03227011859416962
step: 180, loss: 0.00022378721041604877
step: 190, loss: 0.00018443958833813667
/usr/local/lib/python3.7/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=0.7658402203856748, f1=0.7734806629834254,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006789189646951854
step: 10, loss: 0.003954119980335236
step: 20, loss: 0.002147700171917677
step: 30, loss: 0.0013480427442118526
step: 40, loss: 0.000363241444574669
step: 50, loss: 0.08221419155597687
step: 60, loss: 0.0005577327101491392
step: 70, loss: 0.017629366368055344
step: 80, loss: 0.0020790433045476675
step: 90, loss: 0.0012982345651835203
step: 100, loss: 0.0009461325826123357
step: 110, loss: 0.003206223715096712
step: 120, loss: 0.006475933827459812
step: 130, loss: 0.0006770467734895647
step: 140, loss: 0.001614292268641293
step: 150, loss: 0.0004752812092192471
step: 160, loss: 0.033175114542245865
step: 170, loss: 0.0005497073871083558
step: 180, loss: 0.00082878622924909
step: 190, loss: 0.0027709228452295065
/usr/local/lib/python3.7/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=0.7559055118110236, f1=0.7506561679790026,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008938903920352459
step: 10, loss: 0.0005103815929032862
step: 20, loss: 0.0016191533068194985
step: 30, loss: 0.0004567121504805982
step: 40, loss: 0.0017384056700393558
step: 50, loss: 0.0004946588887833059
step: 60, loss: 0.0031859593000262976
step: 70, loss: 0.00036403234116733074
step: 80, loss: 0.0005197254940867424
step: 90, loss: 0.005657298490405083
step: 100, loss: 0.00582754286006093
step: 110, loss: 0.0016514165326952934
step: 120, loss: 0.002087377244606614
step: 130, loss: 0.009376759640872478
step: 140, loss: 0.0004636913363356143
step: 150, loss: 0.0008854361949488521
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.009611563757061958
step: 170, loss: 0.0005688744713552296
step: 180, loss: 0.002132711000740528
step: 190, loss: 0.0022203787229955196
/usr/local/lib/python3.7/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=0.7578947368421052, f1=0.7611548556430445,
best_f1=0.7862796833773087
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0046588340774178505
step: 10, loss: 0.0011993157677352428
step: 20, loss: 0.0009119341848418117
step: 30, loss: 0.0003423503949306905
step: 40, loss: 0.0003679059154819697
step: 50, loss: 0.002914020325988531
step: 60, loss: 0.004946446977555752
step: 70, loss: 0.0022599538788199425
step: 80, loss: 0.0005317020695656538
step: 90, loss: 0.0005011288449168205
step: 100, loss: 0.0003811796777881682
step: 110, loss: 0.0007749670767225325
step: 120, loss: 0.0006490892847068608
step: 130, loss: 0.006494549568742514
step: 140, loss: 0.015641408041119576
step: 150, loss: 0.0012665787944570184
step: 160, loss: 0.00023672683164477348
step: 170, loss: 0.0006722027901560068

```

```

step: 180, loss: 0.0040395366959273815
step: 190, loss: 0.0008492750930599868
/usr/local/lib/python3.7/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=0.7606382978723404, f1=0.7708894878706198,
best_f1=0.7862796833773087
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:08, 233.27it/s]
load_f1 = 0.7087912087912088
real_f1 = 0.6991869918699186
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
4398it [00:16, 267.57it/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: 0.6349101662635803
step: 10, loss: 0.3733975887298584
step: 20, loss: 0.30188921093940735
step: 30, loss: 0.3882068991661072
step: 40, loss: 0.29009875655174255
step: 50, loss: 0.27378326654434204
step: 60, loss: 0.23128829896450043
step: 70, loss: 0.36301106214523315
step: 80, loss: 0.3668130934238434
step: 90, loss: 0.22979453206062317
step: 100, loss: 0.2466941922903061
step: 110, loss: 0.29458439350128174
step: 120, loss: 0.227329820394516
step: 130, loss: 0.10973504185676575
step: 140, loss: 0.21953964233398438
step: 150, loss: 0.2546914517879486
step: 160, loss: 0.1104833260178566
step: 170, loss: 0.2765965759754181
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6901408450704225, f1=0.6711409395973154,
best_f1=0.6711409395973154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1234247237443924
step: 10, loss: 0.13223837316036224
step: 20, loss: 0.10575111210346222
step: 30, loss: 0.19276778399944305
step: 40, loss: 0.06249071657657623
step: 50, loss: 0.10581756383180618
step: 60, loss: 0.052404992282390594
step: 70, loss: 0.07709302008152008
step: 80, loss: 0.10079770535230637
step: 90, loss: 0.07725309580564499
step: 100, loss: 0.1284639835357666
step: 110, loss: 0.02664780057966709
step: 120, loss: 0.10446484386920929
step: 130, loss: 0.041113000363111496
step: 140, loss: 0.21099825203418732
step: 150, loss: 0.228548064827919
step: 160, loss: 0.12087885290384293
step: 170, loss: 0.07453717291355133
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7418546365914787, f1=0.7294117647058824,
best_f1=0.7294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07637009024620056
step: 10, loss: 0.07901676744222641
step: 20, loss: 0.029255203902721405
step: 30, loss: 0.12040424346923828
step: 40, loss: 0.10294388979673386
step: 50, loss: 0.0666002482175827
step: 60, loss: 0.1378566175699234
step: 70, loss: 0.09858869016170502
step: 80, loss: 0.05605373904109001
step: 90, loss: 0.2427312284708023
step: 100, loss: 0.0050373622216284275
step: 110, loss: 0.12021249532699585

```

```

step: 120, loss: 0.06684296578168869
step: 130, loss: 0.219973623752594
step: 140, loss: 0.11188613623380661
step: 150, loss: 0.04779529199004173
step: 160, loss: 0.04561798647046089
step: 170, loss: 0.1459435224533081
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7548209366391185, f1=0.7835051546391752,
best_f1=0.7835051546391752
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04822376370429993
step: 10, loss: 0.14961890876293182
step: 20, loss: 0.029539896175265312
step: 30, loss: 0.025713521987199783
step: 40, loss: 0.015125536359846592
step: 50, loss: 0.04051748663187027
step: 60, loss: 0.08367199450731277
step: 70, loss: 0.01356862485408783
step: 80, loss: 0.03358723968267441
step: 90, loss: 0.06259535998106003
step: 100, loss: 0.1508936882019043
step: 110, loss: 0.09677248448133469
step: 120, loss: 0.04735105112195015
step: 130, loss: 0.03541755676269531
step: 140, loss: 0.022476376965641975
step: 150, loss: 0.06493458151817322
step: 160, loss: 0.12844866514205933
step: 170, loss: 0.11450714617967606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7598039215686274, f1=0.7621247113163971,
best_f1=0.7621247113163971
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```



```

step: 0, loss: 0.07061909884214401
step: 10, loss: 0.12285052239894867
step: 20, loss: 0.054036419838666916
step: 30, loss: 0.15734833478927612
step: 40, loss: 0.016908753663301468
step: 50, loss: 0.03225517272949219
step: 60, loss: 0.045238226652145386
step: 70, loss: 0.14266014099121094
step: 80, loss: 0.06083041429519653
step: 90, loss: 0.10829967260360718
step: 100, loss: 0.01955145224928856
step: 110, loss: 0.07082239538431168
step: 120, loss: 0.01261519268155098
step: 130, loss: 0.04284605011343956
step: 140, loss: 0.07522385567426682
step: 150, loss: 0.08277721703052521
step: 160, loss: 0.1278148591518402
step: 170, loss: 0.00693058967590332
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.797979797979798, f1=0.8083140877598153,
best_f1=0.8083140877598153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009156487882137299
step: 10, loss: 0.01996326632797718
step: 20, loss: 0.010390842333436012
step: 30, loss: 0.013773275539278984
step: 40, loss: 0.011284854263067245
step: 50, loss: 0.14380864799022675
step: 60, loss: 0.1610913872718811
step: 70, loss: 0.0770256295800209
step: 80, loss: 0.06491522490978241
step: 90, loss: 0.04497101902961731
step: 100, loss: 0.008048031479120255
step: 110, loss: 0.0017207748023793101
step: 120, loss: 0.007945179007947445
step: 130, loss: 0.0077005913481116295
step: 140, loss: 0.014789191074669361
step: 150, loss: 0.029643170535564423
step: 160, loss: 0.17914210259914398
step: 170, loss: 0.030829327180981636

```

```

/usr/local/lib/python3.7/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=0.78, f1=0.771689497716895, best_f1=0.8083140877598153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009871645830571651
step: 10, loss: 0.0248886551707983
step: 20, loss: 0.00137730548158288
step: 30, loss: 0.0019323753658682108
step: 40, loss: 0.004837095271795988
step: 50, loss: 0.07047946751117706
step: 60, loss: 0.002754013519734144
step: 70, loss: 0.0009695951011963189
step: 80, loss: 0.008941047824919224
step: 90, loss: 0.006054500583559275
step: 100, loss: 0.006938259117305279
step: 110, loss: 0.011805917136371136
step: 120, loss: 0.004820120055228472
step: 130, loss: 0.1270166039466858
step: 140, loss: 0.013089381158351898
step: 150, loss: 0.0017387360567227006
step: 160, loss: 0.004260243382304907
step: 170, loss: 0.03943953290581703
/usr/local/lib/python3.7/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=0.7906976744186047, f1=0.7924528301886793,
best_f1=0.8083140877598153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021588226780295372
step: 10, loss: 0.0010783844627439976
step: 20, loss: 0.09193672984838486
step: 30, loss: 0.027894018217921257
step: 40, loss: 0.000910491740796715
step: 50, loss: 0.0030819368548691273
step: 60, loss: 0.004133069422096014
step: 70, loss: 0.0010539793875068426
step: 80, loss: 0.005567938555032015
step: 90, loss: 0.007092003244906664
step: 100, loss: 0.0705108642578125
step: 110, loss: 0.013854943215847015
step: 120, loss: 0.021193310618400574
step: 130, loss: 0.009182050824165344
step: 140, loss: 0.001872413675300777

```

```

step: 150, loss: 0.05048566684126854
step: 160, loss: 0.021060774102807045
step: 170, loss: 0.004649720620363951
/usr/local/lib/python3.7/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=0.7844155844155845, f1=0.7895981087470448,
best_f1=0.8083140877598153
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014360406203195453
step: 10, loss: 0.023530248552560806
step: 20, loss: 0.002449322259053588
step: 30, loss: 0.010052538476884365
step: 40, loss: 0.00793358776718378
step: 50, loss: 0.12445204704999924
step: 60, loss: 0.05930840224027634
step: 70, loss: 0.09367312490940094
step: 80, loss: 0.03275679796934128
step: 90, loss: 0.030413281172513962
step: 100, loss: 0.008737563155591488
step: 110, loss: 0.0021033291704952717
step: 120, loss: 0.04805691912770271
step: 130, loss: 0.14062778651714325
step: 140, loss: 0.05504371225833893
step: 150, loss: 0.012118052691221237
step: 160, loss: 0.022405270487070084
step: 170, loss: 0.0024788351729512215
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8098765432098766, f1=0.7857142857142857,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06346803903579712
step: 10, loss: 0.001476424396969378
step: 20, loss: 0.0025778443086892366
step: 30, loss: 0.11623257398605347
step: 40, loss: 0.0007492968579754233
step: 50, loss: 0.06382530182600021
step: 60, loss: 0.0004490888677537441

```

```

step: 70, loss: 0.02231474034488201
step: 80, loss: 0.0006740691023878753
step: 90, loss: 0.01600271463394165
step: 100, loss: 0.0008880191599018872
step: 110, loss: 0.0014237186405807734
step: 120, loss: 0.004763900768011808
step: 130, loss: 0.0034678231459110975
step: 140, loss: 0.12036086618900299
step: 150, loss: 0.08551856130361557
step: 160, loss: 0.004897233098745346
step: 170, loss: 0.001955605112016201
/usr/local/lib/python3.7/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=0.7855421686746988, f1=0.7818181818181817,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03578028455376625
step: 10, loss: 0.005399651359766722
step: 20, loss: 0.006619045976549387
step: 30, loss: 0.0003066071658395231
step: 40, loss: 0.0003075927961617708
step: 50, loss: 0.0039403242990374565
step: 60, loss: 0.007358692120760679
step: 70, loss: 0.00041425187373533845
step: 80, loss: 0.0004068543785251677
step: 90, loss: 0.0007920240750536323
step: 100, loss: 0.0031745859887450933
step: 110, loss: 0.024564528837800026
step: 120, loss: 0.0014736650045961142
step: 130, loss: 0.0012556470464915037
step: 140, loss: 0.030973970890045166
step: 150, loss: 0.01329475175589323
step: 160, loss: 0.00422065332531929
step: 170, loss: 0.0033753577154129744
/usr/local/lib/python3.7/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=0.7729468599033817, f1=0.7755102040816327,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0032984495628625154
step: 10, loss: 0.034098319709300995
step: 20, loss: 0.0015154839493334293

```

```

step: 30, loss: 0.0011925561120733619
step: 40, loss: 0.0006289267330430448
step: 50, loss: 0.000255343911703676
step: 60, loss: 0.0005149513599462807
step: 70, loss: 0.11658070981502533
step: 80, loss: 0.0002257572195958346
step: 90, loss: 0.015386001206934452
step: 100, loss: 0.0007323794416151941
step: 110, loss: 0.0017538216197863221
step: 120, loss: 0.12785407900810242
step: 130, loss: 0.005652558524161577
step: 140, loss: 0.009663957171142101
step: 150, loss: 0.00046781578566879034
step: 160, loss: 0.00040187378181144595
step: 170, loss: 0.0004706756444647908
/usr/local/lib/python3.7/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=0.7611548556430446, f1=0.7823960880195598,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004036376252770424
step: 10, loss: 0.004211035557091236
step: 20, loss: 0.002108511980623007
step: 30, loss: 0.00032747170189395547
step: 40, loss: 0.0010761187877506018
step: 50, loss: 0.0019830227829515934
step: 60, loss: 0.004830663092434406
step: 70, loss: 0.0002833471226040274
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 80, loss: 0.000300778221571818
step: 90, loss: 0.00016268593026325107
step: 100, loss: 0.0001577728835400194
step: 110, loss: 9.878176206257194e-05
step: 120, loss: 0.013093595393002033
step: 130, loss: 0.00014794376329518855
step: 140, loss: 0.0002403700927970931
step: 150, loss: 0.0011167738121002913
step: 160, loss: 9.580074402038008e-05
step: 170, loss: 0.004791234619915485
/usr/local/lib/python3.7/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=0.7604166666666665, f1=0.7893462469733656,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018510087102185935
step: 10, loss: 0.0008286808733828366
step: 20, loss: 0.001182180130854249
step: 30, loss: 0.00030198594322428107
step: 40, loss: 0.0001067103075911291
step: 50, loss: 0.0023117815144360065
step: 60, loss: 8.027419244172052e-05
step: 70, loss: 0.001758192665874958
step: 80, loss: 0.0005280922632664442
step: 90, loss: 0.0005710338009521365
step: 100, loss: 0.01911052130162716
step: 110, loss: 0.00016369775403290987
step: 120, loss: 0.004870882257819176
step: 130, loss: 0.0004528301942627877
step: 140, loss: 0.010790103115141392
step: 150, loss: 7.777903374517336e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.0010944611858576536
step: 170, loss: 0.0002214818960055709
/usr/local/lib/python3.7/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=0.7663043478260869, f1=0.7841191066997518,
best_f1=0.7857142857142857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00022597654606215656
step: 10, loss: 0.002675937954336405
step: 20, loss: 0.006091064307838678
step: 30, loss: 0.000734248198568821
step: 40, loss: 0.00012637648615054786
step: 50, loss: 0.06124657019972801
step: 60, loss: 0.010591704398393631
step: 70, loss: 0.0007568770088255405
step: 80, loss: 0.009095556102693081
step: 90, loss: 0.000649642723146826
step: 100, loss: 0.002303843852132559
step: 110, loss: 0.0014009709702804685
step: 120, loss: 0.0010100622894242406
step: 130, loss: 0.0013971825828775764
step: 140, loss: 0.0013042896753177047
step: 150, loss: 0.00036785233533009887
step: 160, loss: 0.0002768410195130855
step: 170, loss: 0.00013021219638176262
/usr/local/lib/python3.7/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=0.7608695652173912, f1=0.7819548872180452,
best_f1=0.7857142857142857
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 295.91it/s]
load_f1 = 0.4317180616740088
real_f1 = 0.4112676056338029
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
4398it [00:16, 271.75it/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, 498kB/s]

Downloading: 100% 232k/232k [00:00<00:00, 1.21MB/s]

Downloading: 100% 440M/440M [00:43<00:00, 10.2MB/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: 0.632515013217926
step: 10, loss: 0.642586886882782
step: 20, loss: 0.48570528626441956
step: 30, loss: 0.3409290909767151
step: 40, loss: 0.11028208583593369
step: 50, loss: 0.2045281082391739
step: 60, loss: 0.07137104868888855
step: 70, loss: 0.0794663280248642
step: 80, loss: 0.06400313228368759
step: 90, loss: 0.14387458562850952
step: 100, loss: 0.010596612468361855
step: 110, loss: 0.19680488109588623
step: 120, loss: 0.014248291961848736
step: 130, loss: 0.15658271312713623
step: 140, loss: 0.00864881370216608
step: 150, loss: 0.09316825866699219
step: 160, loss: 0.05390497297048569
step: 170, loss: 0.08733851462602615
step: 180, loss: 0.04188191890716553
step: 190, loss: 0.05350013077259064
step: 200, loss: 0.11071658134460449

```

```

step: 210, loss: 0.007079010829329491
step: 220, loss: 0.019978079944849014
step: 230, loss: 0.16389219462871552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9573033707865168, f1=0.9508571428571428,
best_f1=0.9508571428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008707171306014061
step: 10, loss: 0.004445025231689215
step: 20, loss: 0.2732962667942047
step: 30, loss: 0.13314373791217804
step: 40, loss: 0.14821404218673706
step: 50, loss: 0.003306318772956729
step: 60, loss: 0.0044740731827914715
step: 70, loss: 0.12940438091754913
step: 80, loss: 0.005742373410612345
step: 90, loss: 0.04643292352557182
step: 100, loss: 0.03883311524987221
step: 110, loss: 0.06004185974597931
step: 120, loss: 0.048261791467666626
step: 130, loss: 0.034429486840963364
step: 140, loss: 0.001882216427475214
step: 150, loss: 0.011766569688916206
step: 160, loss: 0.040379010140895844
step: 170, loss: 0.0349145270884037
step: 180, loss: 0.05707649886608124
step: 190, loss: 0.05429858714342117
step: 200, loss: 0.0033750608563423157
step: 210, loss: 0.01441223919391632
step: 220, loss: 0.06120113655924797
step: 230, loss: 0.014730937778949738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9612403100775194, f1=0.9620535714285715,
best_f1=0.9620535714285715

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0061707934364676476
step: 10, loss: 0.008716066367924213
step: 20, loss: 0.10662104189395905
step: 30, loss: 0.011383439414203167
step: 40, loss: 0.040450386703014374
step: 50, loss: 0.03347443789243698
step: 60, loss: 0.011155272834002972
step: 70, loss: 0.00203660037368536
step: 80, loss: 0.002772317035123706
step: 90, loss: 0.03283381462097168
step: 100, loss: 0.0007043723599053919
step: 110, loss: 0.001758666941896081
step: 120, loss: 0.006520633120089769
step: 130, loss: 0.001816964359022677
step: 140, loss: 0.008660449646413326
step: 150, loss: 0.010924352332949638
step: 160, loss: 0.09634283185005188
step: 170, loss: 0.005196617916226387
step: 180, loss: 0.008211025968194008
step: 190, loss: 0.013715066947042942
step: 200, loss: 0.15246018767356873
step: 210, loss: 0.019636817276477814
step: 220, loss: 0.03220367431640625
step: 230, loss: 0.0005716328741982579
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9688888888888889, 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: 0.003485091496258974
step: 10, loss: 0.0032758228480815887
step: 20, loss: 0.003009666921570897
step: 30, loss: 0.000701277342159301
step: 40, loss: 0.0012509621446952224
step: 50, loss: 0.00024041836149990559
step: 60, loss: 0.0007630693144164979
step: 70, loss: 0.05427068844437599
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0

```

```

step: 80, loss: 0.0007083013188093901
step: 90, loss: 0.0038908382412046194
step: 100, loss: 0.004336394835263491
step: 110, loss: 0.0012652447912842035
step: 120, loss: 0.030641455203294754
step: 130, loss: 0.00178443337790668
step: 140, loss: 0.0021300893276929855
step: 150, loss: 0.11849034577608109
step: 160, loss: 0.0030662785284221172
step: 170, loss: 0.002556981984525919
step: 180, loss: 0.0009786596056073904
step: 190, loss: 0.0036179644521325827
step: 200, loss: 0.03304629027843475
step: 210, loss: 0.008109665475785732
step: 220, loss: 0.0003810659109149128
step: 230, loss: 0.030986882746219635
/usr/local/lib/python3.7/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=0.9688195991091313, f1=0.971815107102593,
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: 0.0313764363527298
step: 10, loss: 0.00110853579826653
step: 20, loss: 0.03255898505449295
step: 30, loss: 0.0008322803187184036
step: 40, loss: 0.0006541418260894716
step: 50, loss: 0.0004396237200126052
step: 60, loss: 0.0013294850941747427
step: 70, loss: 0.0005191076779738069
step: 80, loss: 0.0003482549509499222
step: 90, loss: 0.000933431088924408
step: 100, loss: 0.00018260625074617565
step: 110, loss: 0.0002487288147676736
step: 120, loss: 0.00017778777692001313
step: 130, loss: 0.00755632109940052
step: 140, loss: 0.02235453389585018
step: 150, loss: 0.008052773773670197
step: 160, loss: 0.004594599828124046
step: 170, loss: 0.0128934346139431
step: 180, loss: 0.011025452986359596
step: 190, loss: 0.07651176303625107
step: 200, loss: 0.0014585979515686631
step: 210, loss: 0.0004641262348741293
step: 220, loss: 0.00484930444508791
step: 230, loss: 0.0011000155936926603

```

```

/usr/local/lib/python3.7/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=0.9663677130044843, f1=0.9628796400449944,
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: 0.003897473681718111
step: 10, loss: 0.007727705407887697
step: 20, loss: 0.002828221768140793
step: 30, loss: 0.002845784416422248
step: 40, loss: 0.0003880493459291756
step: 50, loss: 0.0006998783210292459
step: 60, loss: 0.0003698954242281616
step: 70, loss: 0.0003917161957360804
step: 80, loss: 0.0002928819158114493
step: 90, loss: 0.000907960522454232
step: 100, loss: 0.03363803029060364
step: 110, loss: 0.008714649826288223
step: 120, loss: 0.0006857209373265505
step: 130, loss: 0.00036949425702914596
step: 140, loss: 0.0003002323501277715
step: 150, loss: 0.030111163854599
step: 160, loss: 0.001437175553292036
step: 170, loss: 0.004394026938825846
step: 180, loss: 0.00626629451289773
step: 190, loss: 0.0015360231045633554
step: 200, loss: 0.00023735317518003285
step: 210, loss: 0.0002849874726962298
step: 220, loss: 0.0002037037193076685
step: 230, loss: 0.06575588136911392
/usr/local/lib/python3.7/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=0.966740576496674, f1=0.967525195968645,
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: 0.0008191352826543152
step: 10, loss: 0.0004027953837066889
step: 20, loss: 0.0001653027575230226
step: 30, loss: 0.0010259858099743724
step: 40, loss: 0.0005585222388617694
step: 50, loss: 0.001976566156372428
step: 60, loss: 0.0014712570700794458
step: 70, loss: 0.002641488565132022

```

```

step: 80, loss: 0.0067573427222669125
step: 90, loss: 0.006789698265492916
step: 100, loss: 0.0003935233980882913
step: 110, loss: 0.0017026205314323306
step: 120, loss: 0.0008116101380437613
step: 130, loss: 0.0010968308197334409
step: 140, loss: 0.00027804929413832724
step: 150, loss: 0.010885702446103096
step: 160, loss: 0.03949431702494621
step: 170, loss: 0.003121807938441634
step: 180, loss: 0.003114062361419201
step: 190, loss: 0.0002449059102218598
step: 200, loss: 0.03408607095479965
step: 210, loss: 9.969292295863852e-05
step: 220, loss: 0.00023239814618136734
step: 230, loss: 0.0046981330960989
/usr/local/lib/python3.7/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=0.9626274065685164, f1=0.9627118644067796,
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: 0.00025787318008951843
step: 10, loss: 0.0005329395062290132
step: 20, loss: 0.00016181808314286172
step: 30, loss: 0.0001452345895813778
step: 40, loss: 0.008658148348331451
step: 50, loss: 0.0033689162228256464
step: 60, loss: 9.840037091635168e-05
step: 70, loss: 0.00023010374570731074
step: 80, loss: 0.0002559804415795952
step: 90, loss: 7.615715003339574e-05
step: 100, loss: 0.00015304828411899507
step: 110, loss: 0.00014258756709750742
step: 120, loss: 0.00042053338256664574
step: 130, loss: 0.000342707586241886
step: 140, loss: 5.749555930378847e-05
step: 150, loss: 5.652638719766401e-05
step: 160, loss: 6.097367077018134e-05
step: 170, loss: 0.00010083952656714246
step: 180, loss: 0.00018755563360173255
step: 190, loss: 0.0006454075337387621
step: 200, loss: 0.0001500455109635368
step: 210, loss: 0.00932761188596487
step: 220, loss: 0.0023782984353601933
step: 230, loss: 0.00025820211158134043

```

```

/usr/local/lib/python3.7/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=0.9683972911963882, f1=0.9716231555051079,
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: 9.577813762007281e-05
step: 10, loss: 0.0005313002620823681
step: 20, loss: 0.0012607163516804576
step: 30, loss: 9.089201921597123e-05
step: 40, loss: 0.0022128578275442123
step: 50, loss: 0.0009551187977194786
step: 60, loss: 0.00037732371129095554
step: 70, loss: 0.002642866689711809
step: 80, loss: 0.0006920718587934971
step: 90, loss: 0.0002371771406615153
step: 100, loss: 0.00020843898528255522
step: 110, loss: 6.226902769412845e-05
step: 120, loss: 0.00015422103751916438
step: 130, loss: 0.0003672555321827531
step: 140, loss: 4.2413201299495995e-05
step: 150, loss: 0.00046992284478619695
step: 160, loss: 0.021673260256648064
step: 170, loss: 0.0006622576620429754
step: 180, loss: 0.0006855320534668863
step: 190, loss: 8.93630858627148e-05
step: 200, loss: 0.00017619537538848817
step: 210, loss: 4.9384932935936376e-05
step: 220, loss: 0.0001947819400811568
step: 230, loss: 0.00017853373719844967
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9711111111111111, f1=0.9633740288568259,
best_f1=0.9633740288568259
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.389554957626387e-05
step: 10, loss: 3.3515007089590654e-05
step: 20, loss: 4.001132765552029e-05
step: 30, loss: 0.00010455241135787219

```

```

step: 40, loss: 3.8502985262311995e-05
step: 50, loss: 0.00012185226660221815
step: 60, loss: 0.0002829351869877428
step: 70, loss: 0.00031378972926177084
step: 80, loss: 7.435508450726047e-05
step: 90, loss: 0.0013448643730953336
step: 100, loss: 4.9669590225676075e-05
step: 110, loss: 0.00011564144369913265
step: 120, loss: 5.515383600140922e-05
step: 130, loss: 0.00010495687456568703
step: 140, loss: 0.0420105904340744
step: 150, loss: 0.030273662880063057
step: 160, loss: 3.4210755984531716e-05
step: 170, loss: 4.247954711900093e-05
step: 180, loss: 0.00010681849380489439
step: 190, loss: 0.02281268686056137
step: 200, loss: 3.0434553991653956e-05
step: 210, loss: 7.358835136983544e-05
step: 220, loss: 8.084547880571336e-05
step: 230, loss: 0.00031317787943407893
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9713024282560706, f1=0.966740576496674,
best_f1=0.966740576496674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012603153300005943
step: 10, loss: 0.00012424009037204087
step: 20, loss: 0.002612991724163294
step: 30, loss: 0.0003116040024906397
step: 40, loss: 0.0003168912953697145
step: 50, loss: 0.002869271906092763
step: 60, loss: 0.012627310119569302
step: 70, loss: 0.0012963837943971157
step: 80, loss: 9.745819988893345e-05
step: 90, loss: 0.002131757326424122
step: 100, loss: 0.00011934623762499541
step: 110, loss: 0.00011732470011338592
step: 120, loss: 0.0025926369708031416
step: 130, loss: 3.684769762912765e-05
step: 140, loss: 6.282639515120536e-05
step: 150, loss: 0.0034851953387260437

```



```

step: 160, loss: 9.661618969403207e-05
step: 170, loss: 0.0001241008867509663
step: 180, loss: 0.0001238172990269959
step: 190, loss: 6.446053157560527e-05
step: 200, loss: 0.00026765745133161545
step: 210, loss: 3.214850221411325e-05
step: 220, loss: 5.1684652135008946e-05
step: 230, loss: 8.315704326378182e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9720044792833147, f1=0.9740698985343857,
best_f1=0.9740698985343857
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00012665046961046755
step: 10, loss: 3.84167542506475e-05
step: 20, loss: 6.65660627419129e-05
step: 30, loss: 0.0005343802622519433
step: 40, loss: 0.001211248803883791
step: 50, loss: 0.0006840151036158204
step: 60, loss: 0.0008807973354123533
step: 70, loss: 0.0006700429366901517
step: 80, loss: 0.00010185245628235862
step: 90, loss: 5.44132744835224e-05
step: 100, loss: 3.057963840547018e-05
step: 110, loss: 4.765520498040132e-05
step: 120, loss: 0.00014860588998999447
step: 130, loss: 5.140569555805996e-05
step: 140, loss: 5.210328890825622e-05
step: 150, loss: 9.036024857778102e-05
step: 160, loss: 0.003976966254413128
step: 170, loss: 3.1209456210490316e-05
step: 180, loss: 3.8092650356702507e-05
step: 190, loss: 0.031920842826366425
step: 200, loss: 5.250625326880254e-05
step: 210, loss: 6.776987720513716e-05
step: 220, loss: 0.0006050681113265455
step: 230, loss: 0.00020283233607187867
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9732739420935412, f1=0.9741863075196409,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.142366328276694e-05
step: 10, loss: 9.809080074774101e-05
step: 20, loss: 6.375390512403101e-05
step: 30, loss: 0.00020143736037425697
step: 40, loss: 4.8878406232688576e-05
step: 50, loss: 0.010259118862450123
step: 60, loss: 0.05733020603656769
step: 70, loss: 0.002289160154759884
step: 80, loss: 6.055839185137302e-05
step: 90, loss: 8.843211981002241e-05
step: 100, loss: 3.778716563829221e-05
step: 110, loss: 0.016159774735569954
step: 120, loss: 0.0010747717460617423
step: 130, loss: 5.4898278904147446e-05
step: 140, loss: 0.00011628969514276832
step: 150, loss: 4.0477119910065085e-05
step: 160, loss: 6.0830840084236115e-05
step: 170, loss: 0.0001112866448238492
step: 180, loss: 3.505387212499045e-05
step: 190, loss: 3.563069549272768e-05
step: 200, loss: 7.888636901043355e-05
step: 210, loss: 2.1900526917306706e-05
step: 220, loss: 3.9582911995239556e-05
step: 230, loss: 5.4418585932580754e-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=0.9696969696969697, f1=0.9708520179372198,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002459439856465906
step: 10, loss: 6.914579716976732e-05
step: 20, loss: 2.5975741664296947e-05
step: 30, loss: 4.036188329337165e-05
step: 40, loss: 3.506320354063064e-05
step: 50, loss: 5.294669608701952e-05
step: 60, loss: 0.0008331572753377259
step: 70, loss: 6.642404332524166e-05
```

```

step: 80, loss: 5.59004838578403e-05
step: 90, loss: 8.489149331580848e-05
step: 100, loss: 3.712359830387868e-05
step: 110, loss: 0.0001552381581859663
step: 120, loss: 2.2719978005625308e-05
step: 130, loss: 3.5698350984603167e-05
step: 140, loss: 0.00013094209134578705
step: 150, loss: 2.6858462661039084e-05
step: 160, loss: 0.00012578078894875944
step: 170, loss: 2.2947246179683134e-05
step: 180, loss: 0.0001427863899152726
step: 190, loss: 5.929296821705066e-05
step: 200, loss: 4.7211669880198315e-05
step: 210, loss: 3.528477463987656e-05
step: 220, loss: 7.426505908370018e-05
step: 230, loss: 4.523422830970958e-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=0.9765363128491621, f1=0.9741863075196409,
best_f1=0.9741863075196409
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.7081820007879287e-05
step: 10, loss: 1.985157359740697e-05
step: 20, loss: 4.268203701940365e-05
step: 30, loss: 0.00040701578836888075
step: 40, loss: 0.00225495221093297
step: 50, loss: 0.024871936067938805
step: 60, loss: 0.00010364632908022031
step: 70, loss: 3.9731508877594024e-05
step: 80, loss: 0.0004954891628585756
step: 90, loss: 6.046359703759663e-05
step: 100, loss: 5.593514651991427e-05
step: 110, loss: 3.016206574102398e-05
step: 120, loss: 6.362241401802748e-05
step: 130, loss: 3.712862962856889e-05
step: 140, loss: 2.974550261569675e-05
step: 150, loss: 9.531294927001e-05
step: 160, loss: 3.359527181601152e-05
step: 170, loss: 2.406844578217715e-05
step: 180, loss: 0.00025534952874295413
step: 190, loss: 0.0009832537034526467

```

```

step: 200, loss: 4.2562358430586755e-05
step: 210, loss: 7.216700032586232e-05
step: 220, loss: 5.792981755803339e-05
step: 230, loss: 0.0001079776047845371
/usr/local/lib/python3.7/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=0.9765363128491621, f1=0.9730941704035874,
best_f1=0.9741863075196409
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 [00:12, 194.28it/s]
load_f1 = 0.9723756906077348
real_f1 = 0.9680264608599779
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, 227.21it/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: 0.6217563152313232
step: 10, loss: 0.5185362100601196
step: 20, loss: 0.528659462928772
step: 30, loss: 0.20974169671535492
step: 40, loss: 0.23954269289970398
step: 50, loss: 0.19369393587112427
step: 60, loss: 0.11263235658407211
step: 70, loss: 0.16929379105567932
step: 80, loss: 0.036021456122398376
step: 90, loss: 0.3868634104728699
step: 100, loss: 0.058321814984083176
step: 110, loss: 0.17388930916786194
step: 120, loss: 0.16949546337127686
step: 130, loss: 0.08196000754833221
step: 140, loss: 0.10859055817127228
step: 150, loss: 0.06975255161523819
step: 160, loss: 0.03722602501511574
step: 170, loss: 0.25539934635162354
step: 180, loss: 0.19951701164245605
step: 190, loss: 0.00877998024225235
step: 200, loss: 0.22396545112133026
step: 210, loss: 0.06868281960487366
step: 220, loss: 0.10116655379533768
step: 230, loss: 0.21766002476215363
step: 240, loss: 0.11056974530220032
step: 250, loss: 0.03835686668753624

```

```

step: 260, loss: 0.15007208287715912
step: 270, loss: 0.025766704231500626
step: 280, loss: 0.09257613122463226
step: 290, loss: 0.25040900707244873
step: 300, loss: 0.10394662618637085
step: 310, loss: 0.26456400752067566
step: 320, loss: 0.1837003380060196
step: 330, loss: 0.09954865276813507
step: 340, loss: 0.13997527956962585
step: 350, loss: 0.1885589212179184
step: 360, loss: 0.06489942222833633
step: 370, loss: 0.11780085414648056
step: 380, loss: 0.01993892341852188
step: 390, loss: 0.19154612720012665
step: 400, loss: 0.2775323688983917
step: 410, loss: 0.03784952312707901
step: 420, loss: 0.0956198200583458
step: 430, loss: 0.14993180334568024
step: 440, loss: 0.01927666924893856
step: 450, loss: 0.026170678436756134
step: 460, loss: 0.02658405900001526
step: 470, loss: 0.1304294317960739
step: 480, loss: 0.07926707714796066
step: 490, loss: 0.07495051622390747
step: 500, loss: 0.08639046549797058
step: 510, loss: 0.12532928586006165
step: 520, loss: 0.03802694380283356
step: 530, loss: 0.018738681450486183
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9145259224661373, f1=0.9183202584217812,
best_f1=0.9183202584217812
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1174502745270729
step: 10, loss: 0.11198265105485916
step: 20, loss: 0.03513826057314873
step: 30, loss: 0.03839811310172081
step: 40, loss: 0.07544117420911789
step: 50, loss: 0.1918499767780304
step: 60, loss: 0.05411916971206665
step: 70, loss: 0.11501374840736389

```

step: 80, loss: 0.11373814195394516
step: 90, loss: 0.022605741396546364
step: 100, loss: 0.10951391607522964
step: 110, loss: 0.07377022504806519
step: 120, loss: 0.2617008090019226
step: 130, loss: 0.16877806186676025
step: 140, loss: 0.11270293593406677
step: 150, loss: 0.11861884593963623
step: 160, loss: 0.08828790485858917
step: 170, loss: 0.07434945553541183
step: 180, loss: 0.02767482027411461
step: 190, loss: 0.08144467324018478
step: 200, loss: 0.04010473191738129
step: 210, loss: 0.04425504803657532
step: 220, loss: 0.07749725133180618
step: 230, loss: 0.011293178424239159
step: 240, loss: 0.01196235977113247
step: 250, loss: 0.08577824383974075
step: 260, loss: 0.004623805172741413
step: 270, loss: 0.24940446019172668
step: 280, loss: 0.12377195805311203
step: 290, loss: 0.05874595791101456
step: 300, loss: 0.179429829120636
step: 310, loss: 0.023624107241630554
step: 320, loss: 0.18986548483371735
step: 330, loss: 0.06934400647878647
step: 340, loss: 0.07054231315851212
step: 350, loss: 0.0032262869644910097
step: 360, loss: 0.16680796444416046
step: 370, loss: 0.12653300166130066
step: 380, loss: 0.04076289013028145
step: 390, loss: 0.080286405980587
step: 400, loss: 0.05600016564130783
step: 410, loss: 0.05144456773996353
step: 420, loss: 0.014797277748584747
step: 430, loss: 0.017074773088097572
step: 440, loss: 0.1060648113489151
step: 450, loss: 0.09057100862264633
step: 460, loss: 0.02069653384387493
step: 470, loss: 0.0347505621612072
step: 480, loss: 0.19630517065525055
step: 490, loss: 0.013121750205755234
step: 500, loss: 0.3157868981361389
step: 510, loss: 0.039313312619924545
step: 520, loss: 0.10427746921777725
step: 530, loss: 0.11055080592632294
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.


```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9198350893266148, f1=0.9175824175824175,
best_f1=0.9175824175824175
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14752601087093353
step: 10, loss: 0.04301169514656067
step: 20, loss: 0.14183299243450165
step: 30, loss: 0.07481282949447632
step: 40, loss: 0.08399984985589981
step: 50, loss: 0.28551170229911804
step: 60, loss: 0.09912621229887009
step: 70, loss: 0.005815268028527498
step: 80, loss: 0.0025211747270077467
step: 90, loss: 0.006642411928623915
step: 100, loss: 0.04438835754990578
step: 110, loss: 0.01080573070794344
step: 120, loss: 0.007325942628085613
step: 130, loss: 0.028072042390704155
step: 140, loss: 0.09890975058078766
step: 150, loss: 0.09789669513702393
step: 160, loss: 0.022134840488433838
step: 170, loss: 0.11381950974464417
step: 180, loss: 0.020424241200089455
step: 190, loss: 0.00575905479490757
step: 200, loss: 0.012880099937319756
step: 210, loss: 0.04377387464046478
step: 220, loss: 0.13630685210227966
step: 230, loss: 0.07476348429918289
step: 240, loss: 0.07298135757446289
step: 250, loss: 0.03175550326704979
step: 260, loss: 0.016509834676980972
step: 270, loss: 0.016600582748651505
step: 280, loss: 0.1564587503671646
step: 290, loss: 0.018444547429680824
step: 300, loss: 0.008675392717123032
step: 310, loss: 0.04886860027909279
step: 320, loss: 0.007505382411181927
step: 330, loss: 0.009211632423102856
step: 340, loss: 0.03927475959062576
step: 350, loss: 0.008387796580791473
step: 360, loss: 0.02026875503361225
step: 370, loss: 0.03501884266734123

```

```

step: 380, loss: 0.01608789712190628
step: 390, loss: 0.018634753301739693
step: 400, loss: 0.029223496094346046
step: 410, loss: 0.006154345814138651
step: 420, loss: 0.11110235005617142
step: 430, loss: 0.05328451097011566
step: 440, loss: 0.0373140312731266
step: 450, loss: 0.07822777330875397
step: 460, loss: 0.047632988542318344
step: 470, loss: 0.16104555130004883
step: 480, loss: 0.013920470140874386
step: 490, loss: 0.006044168956577778
step: 500, loss: 0.12293573468923569
step: 510, loss: 0.038633398711681366
step: 520, loss: 0.024924710392951965
step: 530, loss: 0.10297740995883942
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9251950435979808, f1=0.9174646602827177,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011501934379339218
step: 10, loss: 0.006143192760646343
step: 20, loss: 0.011187439784407616
step: 30, loss: 0.006300085689872503
step: 40, loss: 0.10656970739364624
step: 50, loss: 0.011358950287103653
step: 60, loss: 0.02408277802169323
step: 70, loss: 0.03826650232076645
step: 80, loss: 0.0019520242931321263
step: 90, loss: 0.14116743206977844
step: 100, loss: 0.0195135697722435
step: 110, loss: 0.021271368488669395
step: 120, loss: 0.021892907097935677
step: 130, loss: 0.02959187887609005
step: 140, loss: 0.009629689157009125
step: 150, loss: 0.0864962562918663
step: 160, loss: 0.05791965126991272
step: 170, loss: 0.0225607231259346
step: 180, loss: 0.011915854178369045
step: 190, loss: 0.0188397616147995

```

```

step: 200, loss: 0.05696810036897659
step: 210, loss: 0.061169542372226715
step: 220, loss: 0.005242137238383293
step: 230, loss: 0.09382117539644241
step: 240, loss: 0.017387978732585907
step: 250, loss: 0.037084102630615234
step: 260, loss: 0.03155040368437767
step: 270, loss: 0.011743546463549137
step: 280, loss: 0.14774876832962036
step: 290, loss: 0.03353733569383621
step: 300, loss: 0.00048484973376616836
step: 310, loss: 0.00494310911744833
step: 320, loss: 0.05927431955933571
step: 330, loss: 0.13166429102420807
step: 340, loss: 0.19164764881134033
step: 350, loss: 0.047946371138095856
step: 360, loss: 0.06767883151769638
step: 370, loss: 0.04299916699528694
step: 380, loss: 0.02901976741850376
step: 390, loss: 0.009637993760406971
step: 400, loss: 0.0007529295980930328
step: 410, loss: 0.0032056362833827734
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 420, loss: 0.17642678320407867
step: 430, loss: 0.10943920910358429
step: 440, loss: 0.005009010899811983
step: 450, loss: 0.00944541022181511
step: 460, loss: 0.004829959943890572
step: 470, loss: 0.1131267249584198
step: 480, loss: 0.04866577684879303
step: 490, loss: 0.011057766154408455
step: 500, loss: 0.008993654511868954
step: 510, loss: 0.014035049825906754
step: 520, loss: 0.09119591116905212
step: 530, loss: 0.030052676796913147
/usr/local/lib/python3.7/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=0.9189439555349699, f1=0.9191919191919192,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009686539880931377
step: 10, loss: 0.05663127824664116
step: 20, loss: 0.005653551314026117
step: 30, loss: 0.0010664118453860283
step: 40, loss: 0.0034200744703412056

```

step: 50, loss: 0.16538400948047638
step: 60, loss: 0.1539832055568695
step: 70, loss: 0.025186553597450256
step: 80, loss: 0.0033476930111646652
step: 90, loss: 0.0047604721039533615
step: 100, loss: 0.010415850207209587
step: 110, loss: 0.0016223996644839644
step: 120, loss: 0.00136427185498178
step: 130, loss: 0.03846132382750511
step: 140, loss: 0.003900137497112155
step: 150, loss: 0.0010695658857002854
step: 160, loss: 0.0018838171381503344
step: 170, loss: 0.03300345689058304
step: 180, loss: 0.0008358415798284113
step: 190, loss: 0.005972633138298988
step: 200, loss: 0.005494397599250078
step: 210, loss: 0.010531209409236908
step: 220, loss: 0.012712971307337284
step: 230, loss: 0.09708607941865921
step: 240, loss: 0.0009875886607915163
step: 250, loss: 0.037088293582201004
step: 260, loss: 0.005805760156363249
step: 270, loss: 0.003658795729279518
step: 280, loss: 0.003801817772909999
step: 290, loss: 0.12870706617832184
step: 300, loss: 0.00627620005980134
step: 310, loss: 0.0005597452400252223
step: 320, loss: 0.09968040138483047
step: 330, loss: 0.008604477159678936
step: 340, loss: 0.006554785650223494
step: 350, loss: 0.0009778772946447134
step: 360, loss: 0.06218770891427994
step: 370, loss: 0.09722001105546951
step: 380, loss: 0.0017296909354627132
step: 390, loss: 0.0006407503969967365
step: 400, loss: 0.0592154823243618
step: 410, loss: 0.002572778146713972
step: 420, loss: 0.002174091525375843
step: 430, loss: 0.05487305298447609
step: 440, loss: 0.028773557394742966
step: 450, loss: 0.010570275597274303
step: 460, loss: 0.007475643884390593
step: 470, loss: 0.10109750926494598
step: 480, loss: 0.004732798784971237
step: 490, loss: 0.0015453404048457742
step: 500, loss: 0.015867402777075768
step: 510, loss: 0.02921769581735134
step: 520, loss: 0.031727004796266556

```
step: 530, loss: 0.08913427591323853
/usr/local/lib/python3.7/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=0.909437559580553, f1=0.9157196969696969,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025604039430618286
step: 10, loss: 0.004634636454284191
step: 20, loss: 0.09059256315231323
step: 30, loss: 0.0018550417153164744
step: 40, loss: 0.029675807803869247
step: 50, loss: 0.016085924580693245
step: 60, loss: 0.0017454701010137796
step: 70, loss: 0.010218623094260693
step: 80, loss: 0.011358664371073246
step: 90, loss: 0.0047526732087135315
step: 100, loss: 0.013384304009377956
step: 110, loss: 0.01404107827693224
step: 120, loss: 0.001774753793142736
step: 130, loss: 0.0990816205739975
step: 140, loss: 0.004161810036748648
step: 150, loss: 0.09930930286645889
step: 160, loss: 0.003938819281756878
step: 170, loss: 0.020587844774127007
step: 180, loss: 0.0005835209158249199
step: 190, loss: 0.002757801441475749
step: 200, loss: 0.0028648697771131992
step: 210, loss: 0.1098112165927887
step: 220, loss: 0.014929983764886856
step: 230, loss: 0.004218988120555878
step: 240, loss: 0.07088574767112732
step: 250, loss: 0.025131968781352043
step: 260, loss: 0.06634501367807388
step: 270, loss: 0.13532011210918427
step: 280, loss: 0.0020011723972857
step: 290, loss: 0.0008166754269041121
step: 300, loss: 0.010964024811983109
step: 310, loss: 0.008536754176020622
step: 320, loss: 0.0015435630921274424
step: 330, loss: 0.0009292737813666463
step: 340, loss: 0.07518082112073898
step: 350, loss: 0.0022662030532956123
step: 360, loss: 0.00740521214902401
step: 370, loss: 0.011883476749062538
step: 380, loss: 0.004472378175705671
```

```

step: 390, loss: 0.015551092103123665
step: 400, loss: 0.0007772153476253152
step: 410, loss: 0.005458640865981579
step: 420, loss: 0.003814173396676779
step: 430, loss: 0.0014043508563190699
step: 440, loss: 0.006241587456315756
step: 450, loss: 0.0001484852546127513
step: 460, loss: 0.05141366645693779
step: 470, loss: 0.01665542833507061
step: 480, loss: 0.001123068737797439
step: 490, loss: 0.006009036675095558
step: 500, loss: 0.00921922829002142
step: 510, loss: 0.017269229516386986
step: 520, loss: 0.005457589402794838
step: 530, loss: 0.010273654013872147
/usr/local/lib/python3.7/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=0.9149232914923291, f1=0.9167816091954023,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003734482452273369
step: 10, loss: 0.000768407538998872
step: 20, loss: 0.010275071486830711
step: 30, loss: 0.001637682900764048
step: 40, loss: 0.0009067517821677029
step: 50, loss: 0.0695062205195427
step: 60, loss: 0.003255739575251937
step: 70, loss: 0.010675166733562946
step: 80, loss: 0.011289029382169247
step: 90, loss: 0.008016102947294712
step: 100, loss: 8.113391231745481e-05
step: 110, loss: 0.0010983203537762165
step: 120, loss: 0.0010792150860652328
step: 130, loss: 0.00130807317327708
step: 140, loss: 0.00010351962555432692
step: 150, loss: 0.11124274134635925
step: 160, loss: 0.002039629267528653
step: 170, loss: 0.001198777463287115
step: 180, loss: 0.0029244539327919483
step: 190, loss: 0.000742583186365664
step: 200, loss: 0.0011950157349929214
step: 210, loss: 0.00709048705175519
step: 220, loss: 0.00040037903818301857
step: 230, loss: 0.0009652001899667084
step: 240, loss: 0.0002542250440455973

```

```

step: 250, loss: 0.02399519644677639
step: 260, loss: 0.00024430020130239427
step: 270, loss: 0.0005863035912625492
step: 280, loss: 0.00030163463088683784
step: 290, loss: 7.915085006970912e-05
step: 300, loss: 0.0025370996445417404
step: 310, loss: 0.00014316110173240304
step: 320, loss: 8.725855877855793e-05
step: 330, loss: 0.018305039033293724
step: 340, loss: 0.09980391710996628
step: 350, loss: 0.0038321861065924168
step: 360, loss: 0.0007325591868720949
step: 370, loss: 0.00015722482930868864
step: 380, loss: 0.005348801612854004
step: 390, loss: 0.026121627539396286
step: 400, loss: 0.0061453054659068584
step: 410, loss: 0.012065837159752846
step: 420, loss: 0.0036588481161743402
step: 430, loss: 0.001399954897351563
step: 440, loss: 0.027823064476251602
step: 450, loss: 0.00034276340738870203
step: 460, loss: 0.000507975579239428
step: 470, loss: 0.00798630528151989
step: 480, loss: 0.09809070825576782
step: 490, loss: 0.057895079255104065
step: 500, loss: 0.009334301576018333
step: 510, loss: 0.005485821049660444
step: 520, loss: 0.014415014535188675
step: 530, loss: 0.008521633222699165
/usr/local/lib/python3.7/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=0.9182736455463728, f1=0.9176149294492489,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034271632321178913
step: 10, loss: 0.014243828132748604
step: 20, loss: 0.000576638849452138
step: 30, loss: 0.07071506232023239
step: 40, loss: 0.0018624382792040706
step: 50, loss: 0.0024699755012989044
step: 60, loss: 0.014319933019578457
step: 70, loss: 0.0007009622640907764
step: 80, loss: 0.0023836358450353146
step: 90, loss: 0.0014176021795719862
step: 100, loss: 0.002229986246675253

```

```

step: 110, loss: 0.003953293897211552
step: 120, loss: 0.057420555502176285
step: 130, loss: 0.002059730002656579
step: 140, loss: 0.0018884413875639439
step: 150, loss: 0.003950256388634443
step: 160, loss: 0.007761215325444937
step: 170, loss: 0.004137315787374973
step: 180, loss: 0.0007038542535156012
step: 190, loss: 0.00452418252825737
step: 200, loss: 0.0010942245135083795
step: 210, loss: 0.01808985322713852
step: 220, loss: 0.00371425854973495
step: 230, loss: 0.0027921844739466906
step: 240, loss: 0.005178879480808973
step: 250, loss: 0.0010559671791270375
step: 260, loss: 0.0005066763260401785
step: 270, loss: 0.06130650267004967
step: 280, loss: 0.003235835814848542
step: 290, loss: 0.003853185335174203
step: 300, loss: 0.01610385626554489
step: 310, loss: 0.005450564902275801
step: 320, loss: 0.02619277499616146
step: 330, loss: 0.02107091434299946
step: 340, loss: 0.003737544408068061
step: 350, loss: 0.0009989844402298331
step: 360, loss: 0.023610442876815796
step: 370, loss: 0.0011208690702915192
step: 380, loss: 0.001655382220633328
step: 390, loss: 0.12692712247371674
step: 400, loss: 0.00882565975189209
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 410, loss: 0.011370580643415451
step: 420, loss: 0.005850100889801979
step: 430, loss: 0.021229639649391174
step: 440, loss: 0.006460669916123152
step: 450, loss: 0.005369971040636301
step: 460, loss: 0.00030409145983867347
step: 470, loss: 0.0009203821537084877
step: 480, loss: 0.0008511372143402696
step: 490, loss: 0.009816989302635193
step: 500, loss: 0.005991769023239613
step: 510, loss: 0.0010354453697800636
step: 520, loss: 0.0020484908018261194
step: 530, loss: 0.02333933115005493
/usr/local/lib/python3.7/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=0.9227889564810482, f1=0.9239981575310916,

```



```
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004147982224822044
step: 10, loss: 8.837541827233508e-05
step: 20, loss: 0.003781047184020281
step: 30, loss: 0.0011413736501708627
step: 40, loss: 0.02226177230477333
step: 50, loss: 0.004975943360477686
step: 60, loss: 0.006685927975922823
step: 70, loss: 0.0008535824599675834
step: 80, loss: 0.0017192417290061712
step: 90, loss: 0.0001254996459465474
step: 100, loss: 0.005582455080002546
step: 110, loss: 0.005734928417950869
step: 120, loss: 0.0027009393088519573
step: 130, loss: 0.10319715738296509
step: 140, loss: 0.0002029176102951169
step: 150, loss: 0.0001929739082697779
step: 160, loss: 0.0023068315349519253
step: 170, loss: 0.013773113489151001
step: 180, loss: 8.836296910885721e-05
step: 190, loss: 0.0004001839552074671
step: 200, loss: 0.010436399839818478
step: 210, loss: 0.0009182948269881308
step: 220, loss: 0.001189122791402042
step: 230, loss: 0.00986037589609623
step: 240, loss: 0.0024363186676055193
step: 250, loss: 0.001500299316830933
step: 260, loss: 0.06249920278787613
step: 270, loss: 0.00030675798188894987
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 280, loss: 0.0009765855502337217
step: 290, loss: 0.03340183570981026
step: 300, loss: 0.00019416007853578776
step: 310, loss: 0.006105377338826656
step: 320, loss: 0.0004488098784349859
step: 330, loss: 0.00019891810370609164
step: 340, loss: 0.013973968103528023
step: 350, loss: 0.005790213588625193
step: 360, loss: 0.0010698933620005846
step: 370, loss: 0.0004031872085761279
step: 380, loss: 0.0005822023958899081
step: 390, loss: 0.0002632205723784864
step: 400, loss: 0.03889031335711479
step: 410, loss: 0.000251934805419296
step: 420, loss: 0.00042921132990159094
```

```
step: 430, loss: 0.0001331416133325547
step: 440, loss: 0.0018485180335119367
step: 450, loss: 0.0002856035716831684
step: 460, loss: 0.013289774768054485
step: 470, loss: 0.00011462329712230712
step: 480, loss: 6.261152884690091e-05
step: 490, loss: 0.002180794719606638
step: 500, loss: 0.004921550862491131
step: 510, loss: 0.002492288127541542
step: 520, loss: 0.009473070502281189
step: 530, loss: 8.62525703269057e-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=0.9185882352941176, f1=0.9258572752548655,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.074679160723463e-05
step: 10, loss: 0.00025707384338602424
step: 20, loss: 7.938336784718558e-05
step: 30, loss: 0.000472605403047055
step: 40, loss: 0.019649406895041466
step: 50, loss: 0.00010922344517894089
step: 60, loss: 9.829731425270438e-05
step: 70, loss: 8.330936543643475e-05
step: 80, loss: 0.0011018087388947606
step: 90, loss: 0.0011071563931182027
step: 100, loss: 0.00011383829405531287
step: 110, loss: 0.0001703844463918358
step: 120, loss: 0.0004142624093219638
step: 130, loss: 0.00875110737979412
step: 140, loss: 0.0071732401847839355
step: 150, loss: 0.00010226992162643
step: 160, loss: 8.383275417145342e-05
step: 170, loss: 0.00271188304759562
step: 180, loss: 0.023488689213991165
step: 190, loss: 0.003158877370879054
step: 200, loss: 0.03517809510231018
step: 210, loss: 0.0007349832449108362
step: 220, loss: 0.004539196379482746
step: 230, loss: 0.014304596930742264
step: 240, loss: 0.005686427466571331
step: 250, loss: 0.14663074910640717
step: 260, loss: 0.007557201664894819
step: 270, loss: 0.00902665127068758
step: 280, loss: 0.0003810609341599047
```

```

step: 290, loss: 0.0005192285170778632
step: 300, loss: 0.00011492316116346046
step: 310, loss: 0.00013249907351564616
step: 320, loss: 0.0007875696755945683
step: 330, loss: 0.0006643979577347636
step: 340, loss: 0.1762540638446808
step: 350, loss: 0.0012528970837593079
step: 360, loss: 0.00013782858150079846
step: 370, loss: 0.005995402578264475
step: 380, loss: 0.031468529254198074
step: 390, loss: 0.0042780605144798756
step: 400, loss: 0.0028122905641794205
step: 410, loss: 0.002429006388410926
step: 420, loss: 0.012888840399682522
step: 430, loss: 0.0016265141312032938
step: 440, loss: 0.0021976574789732695
step: 450, loss: 0.00019825242634397
step: 460, loss: 0.0010133336763828993
step: 470, loss: 0.00028561236103996634
step: 480, loss: 0.004970952868461609
step: 490, loss: 0.0004987919237464666
step: 500, loss: 0.02453945204615593
step: 510, loss: 0.0008200597949326038
step: 520, loss: 0.0018168180249631405
step: 530, loss: 0.0007548542926087976
/usr/local/lib/python3.7/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=0.9204440333024977, f1=0.9249771271729187,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00019902453641407192
step: 10, loss: 0.00022574390459340066
step: 20, loss: 8.001309470273554e-05
step: 30, loss: 0.0002656195720192045
step: 40, loss: 0.00012115306162741035
step: 50, loss: 0.0015586789231747389
step: 60, loss: 0.0012189324479550123
step: 70, loss: 0.0008638356812298298
step: 80, loss: 0.00018014964007306844
step: 90, loss: 5.986296309856698e-05
step: 100, loss: 0.007591857109218836
step: 110, loss: 0.0010744680184870958
step: 120, loss: 0.0030493815429508686
step: 130, loss: 0.0012174805160611868
step: 140, loss: 0.008169568143785

```

```

step: 150, loss: 0.00016228447202593088
step: 160, loss: 0.00020286475773900747
step: 170, loss: 0.00011223067849641666
step: 180, loss: 0.0923956111073494
step: 190, loss: 0.006320939864963293
step: 200, loss: 0.013907834887504578
step: 210, loss: 0.014747225679457188
step: 220, loss: 0.0011747288517653942
step: 230, loss: 0.009381107985973358
step: 240, loss: 0.0005618251161649823
step: 250, loss: 0.00021443409787025303
step: 260, loss: 5.605124169960618e-05
step: 270, loss: 0.004706528969109058
step: 280, loss: 0.00015732347674202174
step: 290, loss: 0.00034953447175212204
step: 300, loss: 0.0005823380197398365
step: 310, loss: 0.027177808806300163
step: 320, loss: 0.000152356835315004
step: 330, loss: 0.0009752095793373883
step: 340, loss: 0.0018902519950643182
step: 350, loss: 0.0001544550177641213
step: 360, loss: 0.0005371569423004985
step: 370, loss: 0.0013362576719373465
step: 380, loss: 0.0004904234665445983
step: 390, loss: 0.00011896469368366525
step: 400, loss: 0.0002370600268477574
step: 410, loss: 0.020145233720541
step: 420, loss: 0.00022671328042633832
step: 430, loss: 0.00020566358580254018
step: 440, loss: 0.015009026043117046
step: 450, loss: 0.00029859310598112643
step: 460, loss: 0.028303343802690506
step: 470, loss: 0.00018740029190666974
step: 480, loss: 0.002475072629749775
step: 490, loss: 0.00046673277392983437
step: 500, loss: 0.00027461082208901644
step: 510, loss: 0.0005896096699871123
step: 520, loss: 0.00020549212058540434
step: 530, loss: 0.00017228422802872956
/usr/local/lib/python3.7/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=0.9170506912442397, f1=0.9208237986270023,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0035702092573046684

```

step: 10, loss: 7.762904715491459e-05
step: 20, loss: 0.0012714850017800927
step: 30, loss: 0.0012073608813807368
step: 40, loss: 0.0004160538665018976
step: 50, loss: 0.014720690436661243
step: 60, loss: 0.01051345281302929
step: 70, loss: 0.0007411204860545695
step: 80, loss: 0.001558296149596572
step: 90, loss: 0.13285687565803528
step: 100, loss: 0.0017570850905030966
step: 110, loss: 0.00552802998572588
step: 120, loss: 8.450917812297121e-05
step: 130, loss: 0.0001196876764879562
step: 140, loss: 0.11428089439868927
step: 150, loss: 0.00011997255205642432
step: 160, loss: 0.0002240406902274117
step: 170, loss: 0.00011860050290124491
step: 180, loss: 0.0004945110413245857
step: 190, loss: 0.0024173157289624214
step: 200, loss: 0.0008156162803061306
step: 210, loss: 0.02740553952753544
step: 220, loss: 0.00013850731193087995
step: 230, loss: 0.0001365634088870138
step: 240, loss: 0.00031828947248868644
step: 250, loss: 0.0008115139207802713
step: 260, loss: 0.0003638619964476675
step: 270, loss: 8.770491695031524e-05
step: 280, loss: 0.0010467958636581898
step: 290, loss: 0.006542374845594168
step: 300, loss: 0.00014233410183805972
step: 310, loss: 0.0012258798815310001
step: 320, loss: 0.00014734367141500115
step: 330, loss: 0.0001878467737697065
step: 340, loss: 0.0004868896212428808
step: 350, loss: 0.0021698083728551865
step: 360, loss: 0.004899096209555864
step: 370, loss: 0.00029409147100523114
step: 380, loss: 0.0006644375971518457
step: 390, loss: 0.0008773928275331855
step: 400, loss: 0.006665350869297981
step: 410, loss: 0.0021141923498362303
step: 420, loss: 0.004373802803456783
step: 430, loss: 0.0027133708354085684
step: 440, loss: 0.042404089123010635
step: 450, loss: 0.0011881670216098428
step: 460, loss: 0.02181342802941799
step: 470, loss: 0.00015229660493787378
step: 480, loss: 0.012250143103301525

```

step: 490, loss: 0.005198618862777948
step: 500, loss: 0.006084786728024483
step: 510, loss: 0.01095537655055523
step: 520, loss: 0.0002255066647194326
step: 530, loss: 0.024790242314338684
/usr/local/lib/python3.7/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=0.920796665122742, f1=0.925756186984418,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011923894053325057
step: 10, loss: 9.357493399875239e-05
step: 20, loss: 0.004368244204670191
step: 30, loss: 0.006468669045716524
step: 40, loss: 0.0001008523249765858
step: 50, loss: 0.0021011284552514553
step: 60, loss: 0.00020511132606770843
step: 70, loss: 0.0010261539136990905
step: 80, loss: 0.00030074568348936737
step: 90, loss: 0.008163295686244965
step: 100, loss: 0.0008296962478198111
step: 110, loss: 0.0011118859983980656
step: 120, loss: 0.0033047304023057222
step: 130, loss: 6.444181053666398e-05
step: 140, loss: 4.241010174155235e-05
step: 150, loss: 0.0014168473426252604
step: 160, loss: 6.636572652496397e-05
step: 170, loss: 0.0001347253128187731
step: 180, loss: 0.00011373758752597496
step: 190, loss: 0.0010269561316817999
step: 200, loss: 0.0013014592695981264
step: 210, loss: 0.00026274501578882337
step: 220, loss: 0.002895804587751627
step: 230, loss: 0.00015526151400990784
step: 240, loss: 0.0010585533455014229
step: 250, loss: 0.0005322841461747885
step: 260, loss: 0.0011385775869712234
step: 270, loss: 0.012893025763332844
step: 280, loss: 5.023988342145458e-05
step: 290, loss: 0.010342186316847801
step: 300, loss: 0.0006417605327442288
step: 310, loss: 0.0009835645323619246
step: 320, loss: 0.0008149456116370857
step: 330, loss: 0.0006620942731387913
step: 340, loss: 0.0004264703020453453

```

```

step: 350, loss: 0.004954120144248009
step: 360, loss: 0.00011036200157832354
step: 370, loss: 8.282621274702251e-05
step: 380, loss: 0.00011239058221690357
step: 390, loss: 0.002541170222684741
step: 400, loss: 0.0007657908136025071
step: 410, loss: 0.0038978192023932934
step: 420, loss: 9.451190271647647e-05
step: 430, loss: 0.01953800581395626
step: 440, loss: 5.956301902187988e-05
step: 450, loss: 0.00010782812023535371
step: 460, loss: 0.0010219912510365248
step: 470, loss: 0.0003292569308541715
step: 480, loss: 0.0004012601566500962
step: 490, loss: 0.0010376936988905072
step: 500, loss: 0.00024576764553785324
step: 510, loss: 0.000880913226865232
step: 520, loss: 0.00435156375169754
step: 530, loss: 9.089234663406387e-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=0.9191489361702128, f1=0.9203373945641987,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00033970444928854704
step: 10, loss: 7.183759589679539e-05
step: 20, loss: 0.0005866995197720826
step: 30, loss: 5.522717037820257e-05
step: 40, loss: 0.0004769757797475904
step: 50, loss: 0.0035564859863370657
step: 60, loss: 2.6127903765882365e-05
step: 70, loss: 0.0022275669034570456
step: 80, loss: 2.061909981421195e-05
step: 90, loss: 0.000448661798145622
step: 100, loss: 0.0006038465071469545
step: 110, loss: 7.306971383513883e-05
step: 120, loss: 6.006751937093213e-05
step: 130, loss: 0.004100047051906586
step: 140, loss: 0.0003238815988879651
step: 150, loss: 6.813877553213388e-05
step: 160, loss: 0.0002923952997662127
step: 170, loss: 0.0010318078566342592
step: 180, loss: 7.972581806825474e-05
step: 190, loss: 0.00023992233036551625
step: 200, loss: 0.0001275994727620855

```

```

step: 210, loss: 0.0005093562067486346
step: 220, loss: 0.00023561489069834352
step: 230, loss: 0.002370659727603197
step: 240, loss: 0.0006922188331373036
step: 250, loss: 0.00024199405743274838
step: 260, loss: 0.005184466019272804
step: 270, loss: 0.00013202163972891867
step: 280, loss: 7.779520092299208e-05
step: 290, loss: 0.0016236023511737585
step: 300, loss: 0.00048818462528288364
step: 310, loss: 0.005723448004573584
step: 320, loss: 0.0004313979879952967
step: 330, loss: 7.919744530227035e-05
step: 340, loss: 0.00017782991926651448
step: 350, loss: 0.0002295806771144271
step: 360, loss: 0.01210180763155222
step: 370, loss: 6.89472581143491e-05
step: 380, loss: 0.00028211172320879996
step: 390, loss: 0.11877045780420303
step: 400, loss: 0.0005114839877933264
step: 410, loss: 0.00010491681314306334
step: 420, loss: 0.0003252976748626679
step: 430, loss: 0.0005071343621239066
step: 440, loss: 0.0013250703923404217
step: 450, loss: 0.005569308064877987
step: 460, loss: 0.0007125437259674072
step: 470, loss: 0.00831100344657898
step: 480, loss: 0.0006342613487504423
step: 490, loss: 0.0001125643975683488
step: 500, loss: 0.0004426889354363084
step: 510, loss: 0.06589341163635254
step: 520, loss: 0.0006362995482049882
step: 530, loss: 0.005678115412592888
/usr/local/lib/python3.7/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=0.9212962962962964, f1=0.9221611721611722,
best_f1=0.9174646602827177
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.671069372212514e-05
step: 10, loss: 0.0004629641189239919
step: 20, loss: 0.00019593455363065004
step: 30, loss: 0.0007766030030325055
step: 40, loss: 0.09884587675333023
step: 50, loss: 0.00016452015552204102
step: 60, loss: 4.6022749302210286e-05

```


step: 70, loss: 0.000324743683449924
step: 80, loss: 0.00018583814380690455
step: 90, loss: 6.411317735910416e-05
step: 100, loss: 0.0002910315524786711
step: 110, loss: 0.00036470999475568533
step: 120, loss: 0.0012887684861198068
step: 130, loss: 0.011371297761797905
step: 140, loss: 0.010496111586689949
step: 150, loss: 7.922924851300195e-05
step: 160, loss: 0.00019189031445421278
step: 170, loss: 0.0005839613731950521
step: 180, loss: 0.011476105079054832
step: 190, loss: 0.00013994067558087409
step: 200, loss: 0.000385129387723282
step: 210, loss: 0.0002393887407379225
step: 220, loss: 0.00012478831922635436
step: 230, loss: 0.000295048434054479
step: 240, loss: 0.0003802075225394219
step: 250, loss: 8.216577407438308e-05
step: 260, loss: 7.59199247113429e-05
step: 270, loss: 3.0348477594088763e-05
step: 280, loss: 4.5217020669952035e-05
step: 290, loss: 3.169731644447893e-05
step: 300, loss: 5.29974713572301e-05
step: 310, loss: 0.001330302911810577
step: 320, loss: 0.00026951791369356215
step: 330, loss: 7.983206160133705e-05
step: 340, loss: 0.000162066615303047
step: 350, loss: 0.00010463589569553733
step: 360, loss: 0.0005968224722892046
step: 370, loss: 0.0013611163012683392
step: 380, loss: 0.00021938139980193228
step: 390, loss: 0.0006468724459409714
step: 400, loss: 5.7332021242473274e-05
step: 410, loss: 0.000546337862033397
step: 420, loss: 0.00028231466421857476
step: 430, loss: 0.00010717361874412745
step: 440, loss: 0.13805480301380157
step: 450, loss: 5.29938806721475e-05
step: 460, loss: 5.9068195696454495e-05
step: 470, loss: 0.0011957735987380147
step: 480, loss: 4.976085983798839e-05
step: 490, loss: 0.1155642569065094
step: 500, loss: 7.337704300880432e-05
step: 510, loss: 0.003429833799600601
step: 520, loss: 3.540758916642517e-05
step: 530, loss: 0.00013174532796256244
/usr/local/lib/python3.7/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=0.9216241737488197, f1=0.9263746505125815,
best_f1=0.9174646602827177
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
5742it [00:24, 232.17it/s]
load_f1 = 0.9234382339126351
real_f1 = 0.9217065166432256
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, 240.47it/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: 0.5754514336585999
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.41666666666666663, f1=0.5, best_f1=0.5
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4815097451210022
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5789473684210527, f1=0.4090909090909091,
best_f1=0.4090909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4098399877548218
/usr/local/lib/python3.7/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=0.5365853658536585, f1=0.43902439024390244,
best_f1=0.4090909090909091
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3025938868522644
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5945945945945946, f1=0.43902439024390244,
best_f1=0.43902439024390244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21911458671092987
/usr/local/lib/python3.7/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=0.5714285714285714, f1=0.56, best_f1=0.43902439024390244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.38353821635246277
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.689655172413793, f1=0.64, best_f1=0.64
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11779070645570755
/usr/local/lib/python3.7/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=0.6486486486486486, f1=0.43902439024390244, best_f1=0.64

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17404553294181824
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.5714285714285714, best_f1=0.64
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0044101811945438385
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.75, f1=0.5806451612903226, best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.024201102554798126
/usr/local/lib/python3.7/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=0.75, f1=0.6206896551724138, best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006945349276065826
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.6470588235294117,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009030059911310673
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.6470588235294117,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016922518610954285
/usr/local/lib/python3.7/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=0.7096774193548386, f1=0.6666666666666666,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.011369781568646431
/usr/local/lib/python3.7/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=0.7096774193548386, f1=0.6666666666666666,
best_f1=0.5806451612903226
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02235411666336754
/usr/local/lib/python3.7/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=0.7096774193548386, f1=0.6666666666666666,
best_f1=0.5806451612903226
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, 92528.89it/s]
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
load_f1 = 0.7096774193548386
real_f1 = 0.75
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, 270.85it/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 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: 0.6129496693611145
step: 10, loss: 0.6375404000282288
step: 20, loss: 0.4277620017528534
step: 30, loss: 0.28810226917266846
step: 40, loss: 0.26520270109176636
step: 50, loss: 0.09274791926145554
step: 60, loss: 0.05461683124303818
step: 70, loss: 0.06618402898311615
step: 80, loss: 0.06476642936468124
step: 90, loss: 0.07131904363632202
step: 100, loss: 0.017403025180101395
step: 110, loss: 0.07739272713661194
step: 120, loss: 0.039883796125650406
```

```

step: 130, loss: 0.0036890546325594187
step: 140, loss: 0.003138255560770631
step: 150, loss: 0.10137316584587097
step: 160, loss: 0.021350713446736336
step: 170, loss: 0.04792998731136322
step: 180, loss: 0.07295144349336624
step: 190, loss: 0.002142801182344556
step: 200, loss: 0.006815146189182997
step: 210, loss: 0.006028422154486179
step: 220, loss: 0.04361306503415108
step: 230, loss: 0.0431070551276207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9694915254237287, f1=0.9580973952434882,
best_f1=0.9580973952434882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.021361857652664185
step: 10, loss: 0.0022264611907303333
step: 20, loss: 0.10712692141532898
step: 30, loss: 0.19790019094944
step: 40, loss: 0.023965388536453247
step: 50, loss: 0.0045559825375676155
step: 60, loss: 0.0013913246802985668
step: 70, loss: 0.22097283601760864
step: 80, loss: 0.038217995315790176
step: 90, loss: 0.07921551167964935
step: 100, loss: 0.042564719915390015
step: 110, loss: 0.08459364622831345
step: 120, loss: 0.044662538915872574
step: 130, loss: 0.0254659540951252
step: 140, loss: 0.04900424927473068
step: 150, loss: 0.017383446916937828
step: 160, loss: 0.04134754464030266
step: 170, loss: 0.009442275390028954
step: 180, loss: 0.004709662403911352
step: 190, loss: 0.07769103348255157
step: 200, loss: 0.0033877568785101175
step: 210, loss: 0.0039790901355445385
step: 220, loss: 0.009939252398908138
step: 230, loss: 0.01182866282761097
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9764309764309763, f1=0.9653631284916202,
best_f1=0.9653631284916202
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0339583158493042
step: 10, loss: 0.0034414336550980806
step: 20, loss: 0.0012856874382123351
step: 30, loss: 0.05162648484110832
step: 40, loss: 0.10526429861783981
step: 50, loss: 0.00803305208683014
step: 60, loss: 0.0021752063184976578
step: 70, loss: 0.008346067741513252
step: 80, loss: 0.0006439880817197263
step: 90, loss: 0.09038043767213821
step: 100, loss: 0.0030139253940433264
step: 110, loss: 0.0007677703979425132
step: 120, loss: 0.008278748020529747
step: 130, loss: 0.03155673295259476
step: 140, loss: 0.012666665017604828
step: 150, loss: 0.0012226506369188428
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 160, loss: 0.017286764457821846
step: 170, loss: 0.0032086381688714027
step: 180, loss: 0.09604720026254654
step: 190, loss: 0.005430400837212801
step: 200, loss: 0.01985541358590126
step: 210, loss: 0.003864741651341319
step: 220, loss: 0.0010857549495995045
step: 230, loss: 0.0007396753062494099
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9910112359550561, f1=0.9743016759776536,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.0015744275879114866
step: 10, loss: 0.0006542433402501047
step: 20, loss: 0.0007484295638278127
step: 30, loss: 0.001127365743741393
step: 40, loss: 0.009473512880504131
step: 50, loss: 0.021080520004034042
step: 60, loss: 0.0010970542207360268
step: 70, loss: 0.0011028676526620984
step: 80, loss: 0.0007992906612344086
step: 90, loss: 0.0012987249065190554
step: 100, loss: 0.0005873513291589916
step: 110, loss: 0.0002498249523341656
step: 120, loss: 0.06147861108183861
step: 130, loss: 0.006372969131916761
step: 140, loss: 0.0004138632502872497
step: 150, loss: 0.15383322536945343
step: 160, loss: 0.0009255342301912606
step: 170, loss: 0.0025441087782382965
step: 180, loss: 0.000403142417781055
step: 190, loss: 0.00041228061309084296
step: 200, loss: 0.0009020318975672126
step: 210, loss: 0.06681688129901886
step: 220, loss: 0.00014427992573473603
step: 230, loss: 0.003459481755271554
/usr/local/lib/python3.7/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=0.9818181818181818, f1=0.9830124575311437,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013353998074308038
step: 10, loss: 0.00044223887380212545
step: 20, loss: 0.0003624427190516144
step: 30, loss: 8.851233724271879e-05
step: 40, loss: 0.0004478997434489429
step: 50, loss: 0.00035274887341074646
step: 60, loss: 0.0017034201882779598
step: 70, loss: 0.00018936069682240486
step: 80, loss: 0.00017381134966854006
step: 90, loss: 0.0002782318042591214
step: 100, loss: 0.0005703953793272376
step: 110, loss: 0.003589578904211521
step: 120, loss: 0.00012165392399765551
step: 130, loss: 0.00391353340819478
step: 140, loss: 0.013741250149905682
step: 150, loss: 0.0027225418016314507

```

```

step: 160, loss: 0.0020522032864391804
step: 170, loss: 0.003186500631272793
step: 180, loss: 0.005268519744277
step: 190, loss: 0.004702058155089617
step: 200, loss: 0.0013064551167190075
step: 210, loss: 0.0002241163165308535
step: 220, loss: 0.0003727532457560301
step: 230, loss: 0.00015070743393152952
/usr/local/lib/python3.7/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=0.9753363228699552,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005849949666298926
step: 10, loss: 0.00023262637841980904
step: 20, loss: 0.000892729323823005
step: 30, loss: 0.0003658281930256635
step: 40, loss: 0.00033886998426169157
step: 50, loss: 0.00017927755834534764
step: 60, loss: 9.711753955343738e-05
step: 70, loss: 0.009340344928205013
step: 80, loss: 0.00084225571481511
step: 90, loss: 0.0006317587685771286
step: 100, loss: 0.022898074239492416
step: 110, loss: 0.0006622656947001815
step: 120, loss: 0.0005280671757645905
step: 130, loss: 0.0013962909579277039
step: 140, loss: 0.00017539468535687774
step: 150, loss: 0.006595706567168236
step: 160, loss: 0.008880693465471268
step: 170, loss: 0.0028491169214248657
step: 180, loss: 0.00197820202447474
step: 190, loss: 0.005365481600165367
step: 200, loss: 0.00862516462802887
step: 210, loss: 0.0014486489817500114
step: 220, loss: 0.0025745683815330267
step: 230, loss: 0.03561199828982353
/usr/local/lib/python3.7/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=0.9819819819819819, f1=0.9797752808988766,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00236222380772233
step: 10, loss: 0.00021380270482040942
step: 20, loss: 0.00042798288632184267
step: 30, loss: 0.00016401908942498267
step: 40, loss: 0.0012161097256466746
step: 50, loss: 0.0003731815377250314
step: 60, loss: 0.007742900401353836
step: 70, loss: 0.0002601500600576401
step: 80, loss: 0.0002714334987103939
step: 90, loss: 9.254673204850405e-05
step: 100, loss: 0.0001730183867039159
step: 110, loss: 0.00020457126083783805
step: 120, loss: 5.937298192293383e-05
step: 130, loss: 0.00019056166638620198
step: 140, loss: 0.00020829844288527966
step: 150, loss: 0.0005826108390465379
step: 160, loss: 0.04037719964981079
step: 170, loss: 0.005856754258275032
step: 180, loss: 0.001398024964146316
step: 190, loss: 0.0001268013147637248
step: 200, loss: 0.023443050682544708
step: 210, loss: 7.288032793439925e-05
step: 220, loss: 0.10108956694602966
step: 230, loss: 0.03804461285471916
/usr/local/lib/python3.7/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=0.9798206278026906, f1=0.9775280898876404,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006884881877340376
step: 10, loss: 0.002170105930417776
step: 20, loss: 0.007650247309356928
step: 30, loss: 0.006039801519364119
step: 40, loss: 0.0017760028131306171
step: 50, loss: 0.0010803454788401723
step: 60, loss: 0.0013416099827736616
step: 70, loss: 0.00034157244954258204
step: 80, loss: 0.0086625711992383
step: 90, loss: 0.0021608974784612656
step: 100, loss: 0.0013418191811069846
step: 110, loss: 0.0008643389446660876
step: 120, loss: 0.0013577472418546677
step: 130, loss: 0.0007511871517635882
step: 140, loss: 0.00011356112372595817
step: 150, loss: 9.290916204918176e-05

```

```

step: 160, loss: 0.00029169811750762165
step: 170, loss: 7.637299131602049e-05
step: 180, loss: 0.00011685190111165866
step: 190, loss: 0.00014670102973468602
step: 200, loss: 6.965562351979315e-05
step: 210, loss: 0.00013156404020264745
step: 220, loss: 0.00013710437633562833
step: 230, loss: 0.00011206539056729525
/usr/local/lib/python3.7/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=0.983050847457627, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020450558804441243
step: 10, loss: 0.00010390285751782358
step: 20, loss: 0.00025301415007561445
step: 30, loss: 7.194047066150233e-05
step: 40, loss: 0.0001999959204113111
step: 50, loss: 4.6649300202261657e-05
step: 60, loss: 0.0010205814614892006
step: 70, loss: 0.00013559315993916243
step: 80, loss: 0.00011777831241488457
step: 90, loss: 0.00017869099974632263
step: 100, loss: 0.00015417550457641482
step: 110, loss: 4.2928295442834496e-05
step: 120, loss: 0.0027995642740279436
step: 130, loss: 0.00015264379908330739
step: 140, loss: 5.717584281228483e-05
step: 150, loss: 0.0005587390623986721
step: 160, loss: 5.857647192897275e-05
step: 170, loss: 0.00016440835315734148
step: 180, loss: 0.0015158579917624593
step: 190, loss: 9.759235399542376e-05
step: 200, loss: 7.068775448715314e-05
step: 210, loss: 0.0007479332271032035
step: 220, loss: 0.0003923253680113703
step: 230, loss: 0.00022834203264210373
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 9: dev_f1=0.9831271091113611, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00021070205548312515
step: 10, loss: 3.9925424061948434e-05
step: 20, loss: 0.00023556704400107265
step: 30, loss: 0.00017521395056974143
step: 40, loss: 6.723563274135813e-05
step: 50, loss: 5.4234504204941913e-05
step: 60, loss: 0.00011758752953028306
step: 70, loss: 0.00310001103207469
step: 80, loss: 6.148249667603523e-05
step: 90, loss: 0.00036010792246088386
step: 100, loss: 8.037850784603506e-05
step: 110, loss: 7.045917300274596e-05
step: 120, loss: 0.000709695159457624
step: 130, loss: 5.85975285503082e-05
step: 140, loss: 0.00011066385923186317
step: 150, loss: 0.0028779353015124798
step: 160, loss: 4.0395709220319986e-05
step: 170, loss: 5.583345046034083e-05
step: 180, loss: 9.536058496451005e-05
step: 190, loss: 0.00022295815870165825
step: 200, loss: 8.015503408387303e-05
step: 210, loss: 0.16722464561462402
step: 220, loss: 0.000369061017408967
step: 230, loss: 0.0014008956495672464
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9808773903262092, f1=0.9807909604519773,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.076155791059136e-05
step: 10, loss: 0.0002787252014968544
step: 20, loss: 8.160428114933893e-05
step: 30, loss: 0.0015460134018212557
step: 40, loss: 9.91354972939007e-05
step: 50, loss: 7.701127469772473e-05
step: 60, loss: 0.00010074477904709056
step: 70, loss: 0.00028279656544327736
step: 80, loss: 4.890447235084139e-05
step: 90, loss: 7.181249384302646e-05
step: 100, loss: 7.620918768225238e-05
step: 110, loss: 0.0002261526242364198
step: 120, loss: 4.57028036180418e-05
step: 130, loss: 4.237044777255505e-05
step: 140, loss: 0.0002225003991043195
step: 150, loss: 0.00011933943460462615

```



```

step: 160, loss: 6.500444578705356e-05
step: 170, loss: 6.317361112451181e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 180, loss: 8.949101902544498e-05
step: 190, loss: 4.248804543749429e-05
step: 200, loss: 9.17277138796635e-05
step: 210, loss: 4.386251384858042e-05
step: 220, loss: 5.816786506329663e-05
step: 230, loss: 8.30954231787473e-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=0.9820627802690582, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.6451936213998124e-05
step: 10, loss: 4.47759339294862e-05
step: 20, loss: 7.035469025140628e-05
step: 30, loss: 5.443822738016024e-05
step: 40, loss: 0.0024336192291229963
step: 50, loss: 9.117367153521627e-05
step: 60, loss: 9.05894412426278e-05
step: 70, loss: 0.00010771643428597599
step: 80, loss: 5.9257068642182276e-05
step: 90, loss: 3.986190495197661e-05
step: 100, loss: 4.2270934500265867e-05
step: 110, loss: 4.33935783803463e-05
step: 120, loss: 7.582946273032576e-05
step: 130, loss: 8.4918639913667e-05
step: 140, loss: 7.657882088096812e-05
step: 150, loss: 6.699138612020761e-05
step: 160, loss: 0.0015615207375958562
step: 170, loss: 4.518541754805483e-05
step: 180, loss: 0.0001225386222358793
step: 190, loss: 4.00744829676114e-05
step: 200, loss: 3.3350825106026605e-05
step: 210, loss: 0.00010294596722815186
step: 220, loss: 8.657936268718913e-05
step: 230, loss: 0.03214907646179199
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 12: dev_f1=0.9842696629213483, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.754795322194695e-05
step: 10, loss: 5.651185711030848e-05
step: 20, loss: 0.00015075437841005623
step: 30, loss: 8.224845805671066e-05
step: 40, loss: 3.992931669927202e-05
step: 50, loss: 5.065257937530987e-05
step: 60, loss: 5.895052890991792e-05
step: 70, loss: 3.467334317974746e-05
step: 80, loss: 3.793583164224401e-05
step: 90, loss: 4.131788227823563e-05
step: 100, loss: 3.5063945688307285e-05
step: 110, loss: 9.883807797450572e-05
step: 120, loss: 0.00010381525498814881
step: 130, loss: 4.993127367924899e-05
step: 140, loss: 5.176833292352967e-05
step: 150, loss: 3.73967232007999e-05
step: 160, loss: 3.317206574138254e-05
step: 170, loss: 7.318529969779775e-05
step: 180, loss: 4.4190281187184155e-05
step: 190, loss: 2.7447189495433122e-05
step: 200, loss: 5.8454417739994824e-05
step: 210, loss: 2.1665746317012236e-05
step: 220, loss: 4.2148611100856215e-05
step: 230, loss: 3.8084072002675384e-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=0.9830890642615557, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.158986484981142e-05
step: 10, loss: 4.2920779378619045e-05
step: 20, loss: 4.4844928197562695e-05
step: 30, loss: 7.833864219719544e-05
step: 40, loss: 9.334102651337162e-05
step: 50, loss: 5.324737139744684e-05
step: 60, loss: 4.281894143787213e-05
step: 70, loss: 7.545825064880773e-05
step: 80, loss: 3.7970276025589556e-05
step: 90, loss: 0.00011406916019041091
step: 100, loss: 0.00017818310880102217
step: 110, loss: 3.0479235647362657e-05
step: 120, loss: 2.136016519216355e-05
step: 130, loss: 4.7204113798215985e-05
step: 140, loss: 5.1577251724665985e-05

```

```

step: 150, loss: 4.9599188059801236e-05
step: 160, loss: 0.000397457362851128
step: 170, loss: 1.7609256246942095e-05
step: 180, loss: 5.149143180460669e-05
step: 190, loss: 4.4900578359374776e-05
step: 200, loss: 2.993891939695459e-05
step: 210, loss: 3.953485065721907e-05
step: 220, loss: 2.3412594600813463e-05
step: 230, loss: 5.31346449861303e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 14: dev_f1=0.9830890642615557, f1=0.9819413092550789,
best_f1=0.9743016759776536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010655359801603481
step: 10, loss: 2.0998657419113442e-05
step: 20, loss: 5.4887274018255994e-05
step: 30, loss: 6.632118311244994e-05
step: 40, loss: 5.0569586164783686e-05
step: 50, loss: 6.0312428104225546e-05
step: 60, loss: 7.021619967417791e-05
step: 70, loss: 3.55296979250852e-05
step: 80, loss: 0.0001382539776386693
step: 90, loss: 4.87578145111911e-05
step: 100, loss: 3.796192686422728e-05
step: 110, loss: 3.5450735595077276e-05
step: 120, loss: 5.4440737585537136e-05
step: 130, loss: 9.0947920398321e-05
step: 140, loss: 2.8087566533940844e-05
step: 150, loss: 0.00010695866512833163
step: 160, loss: 3.279158772784285e-05
step: 170, loss: 2.9558677852037363e-05
step: 180, loss: 7.773865945637226e-05
step: 190, loss: 4.766130950883962e-05
step: 200, loss: 3.126524461549707e-05
step: 210, loss: 3.3492615330033004e-05
step: 220, loss: 5.0559730880195275e-05
step: 230, loss: 3.7917539884801954e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 15: dev_f1=0.9830890642615557, f1=0.9819413092550789,
best_f1=0.9743016759776536
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 [00:11, 219.47it/s]

load_f1 = 0.9910112359550561

real_f1 = 0.990990990990991

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

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

4398it [00:16, 267.07it/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: 0.6401656270027161

step: 10, loss: 0.548403263092041

step: 20, loss: 0.5578916668891907

step: 30, loss: 0.2820023000240326

step: 40, loss: 0.1763148009777069

step: 50, loss: 0.2568945288658142

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

step: 60, loss: 0.12546811997890472

step: 70, loss: 0.21059413254261017

step: 80, loss: 0.0928536057472229

step: 90, loss: 0.41654831171035767

step: 100, loss: 0.12325279414653778

step: 110, loss: 0.19390611350536346

step: 120, loss: 0.15825526416301727

step: 130, loss: 0.051922086626291275

step: 140, loss: 0.057387180626392365

step: 150, loss: 0.07771595567464828

step: 160, loss: 0.04156766086816788

step: 170, loss: 0.22924701869487762

step: 180, loss: 0.1157885417342186

step: 190, loss: 0.011445362120866776

step: 200, loss: 0.15461516380310059

step: 210, loss: 0.09560216963291168

step: 220, loss: 0.14395104348659515

step: 230, loss: 0.1446991264820099

step: 240, loss: 0.1212119534611702

step: 250, loss: 0.045684050768613815

step: 260, loss: 0.16396602988243103

step: 270, loss: 0.02199603244662285

step: 280, loss: 0.04839450120925903

step: 290, loss: 0.22253724932670593

step: 300, loss: 0.09284356236457825

step: 310, loss: 0.19159989058971405

step: 320, loss: 0.108310267329216

step: 330, loss: 0.05126103386282921

step: 340, loss: 0.1055629551410675

step: 350, loss: 0.06981503218412399

step: 360, loss: 0.09868897497653961

```

step: 370, loss: 0.10287521779537201
step: 380, loss: 0.048004016280174255
step: 390, loss: 0.19773609936237335
step: 400, loss: 0.23057270050048828
step: 410, loss: 0.07325434684753418
step: 420, loss: 0.08507898449897766
step: 430, loss: 0.21569636464118958
step: 440, loss: 0.04574543237686157
step: 450, loss: 0.034309159964323044
step: 460, loss: 0.08705279231071472
step: 470, loss: 0.1440696120262146
step: 480, loss: 0.12845635414123535
step: 490, loss: 0.13556265830993652
step: 500, loss: 0.09788487106561661
step: 510, loss: 0.03229653090238571
step: 520, loss: 0.030654048547148705
step: 530, loss: 0.009390585124492645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9201307800093415, f1=0.9094304388422036,
best_f1=0.9094304388422036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11799515783786774
step: 10, loss: 0.06421452760696411
step: 20, loss: 0.02518618106842041
step: 30, loss: 0.05513453856110573
step: 40, loss: 0.038343220949172974
step: 50, loss: 0.1948377937078476
step: 60, loss: 0.039441756904125214
step: 70, loss: 0.04744437336921692
step: 80, loss: 0.04039846360683441
step: 90, loss: 0.017050784081220627
step: 100, loss: 0.041043784469366074
step: 110, loss: 0.10620772838592529
step: 120, loss: 0.10146015882492065
step: 130, loss: 0.05336228758096695
step: 140, loss: 0.07658109068870544
step: 150, loss: 0.04143499955534935
step: 160, loss: 0.0133061483502388
step: 170, loss: 0.06880736351013184
step: 180, loss: 0.03719785809516907

```

```

step: 190, loss: 0.0928630381822586
step: 200, loss: 0.021252162754535675
step: 210, loss: 0.12258660793304443
step: 220, loss: 0.10722000896930695
step: 230, loss: 0.021698186174035072
step: 240, loss: 0.18800708651542664
step: 250, loss: 0.02133638598024845
step: 260, loss: 0.011286098510026932
step: 270, loss: 0.27874770760536194
step: 280, loss: 0.05875053629279137
step: 290, loss: 0.03374618664383888
step: 300, loss: 0.18674324452877045
step: 310, loss: 0.01564345695078373
step: 320, loss: 0.09280672669410706
step: 330, loss: 0.07849274575710297
step: 340, loss: 0.02523350715637207
step: 350, loss: 0.005467272363603115
step: 360, loss: 0.10945781320333481
step: 370, loss: 0.17246553301811218
step: 380, loss: 0.04794597998261452
step: 390, loss: 0.10619291663169861
step: 400, loss: 0.07227381318807602
step: 410, loss: 0.025868268683552742
step: 420, loss: 0.022689884528517723
step: 430, loss: 0.013823182322084904
step: 440, loss: 0.02818380668759346
step: 450, loss: 0.01954442821443081
step: 460, loss: 0.11699336767196655
step: 470, loss: 0.05804017558693886
step: 480, loss: 0.2952296733856201
step: 490, loss: 0.022371530532836914
step: 500, loss: 0.4020850658416748
step: 510, loss: 0.06131164729595184
step: 520, loss: 0.09585634618997574
step: 530, loss: 0.0660601407289505
/usr/local/lib/python3.7/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=0.9140027014858171, f1=0.912375790424571,
best_f1=0.9094304388422036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.28637149930000305
step: 10, loss: 0.0534142442047596
step: 20, loss: 0.08855294436216354
step: 30, loss: 0.2279268056154251
step: 40, loss: 0.024940699338912964

```


step: 50, loss: 0.018772747367620468
step: 60, loss: 0.04428652673959732
step: 70, loss: 0.0060815829783678055
step: 80, loss: 0.0015911273658275604
step: 90, loss: 0.004826363641768694
step: 100, loss: 0.1027444452047348
step: 110, loss: 0.026483887806534767
step: 120, loss: 0.006823614239692688
step: 130, loss: 0.00610200222581625
step: 140, loss: 0.025945499539375305
step: 150, loss: 0.024450326338410378
step: 160, loss: 0.00908638909459114
step: 170, loss: 0.10300327092409134
step: 180, loss: 0.13360126316547394
step: 190, loss: 0.005314178299158812
step: 200, loss: 0.02317933924496174
step: 210, loss: 0.031215069815516472
step: 220, loss: 0.16607868671417236
step: 230, loss: 0.1419452726840973
step: 240, loss: 0.00693834712728858
step: 250, loss: 0.050619449466466904
step: 260, loss: 0.025243712589144707
step: 270, loss: 0.026112401857972145
step: 280, loss: 0.027881793677806854
step: 290, loss: 0.009629840962588787
step: 300, loss: 0.07862608879804611
step: 310, loss: 0.016031835228204727
step: 320, loss: 0.046212922781705856
step: 330, loss: 0.009331751614809036
step: 340, loss: 0.10184039175510406
step: 350, loss: 0.04409712180495262
step: 360, loss: 0.06560022383928299
step: 370, loss: 0.034785155206918716
step: 380, loss: 0.04946495592594147
step: 390, loss: 0.010645432397723198
step: 400, loss: 0.020098671317100525
step: 410, loss: 0.018630104139447212
step: 420, loss: 0.2208569198846817
step: 430, loss: 0.02200811356306076
step: 440, loss: 0.003585233585909009
step: 450, loss: 0.07534842938184738
step: 460, loss: 0.05609137564897537
step: 470, loss: 0.10103005915880203
step: 480, loss: 0.009879985824227333
step: 490, loss: 0.008819356560707092
step: 500, loss: 0.0063043925911188126
step: 510, loss: 0.018696192651987076
step: 520, loss: 0.036079879850149155

```
step: 530, loss: 0.058836039155721664
/usr/local/lib/python3.7/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=0.9145454545454544, f1=0.9166666666666667,
best_f1=0.9094304388422036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03262300416827202
step: 10, loss: 0.008293005637824535
step: 20, loss: 0.0015189124969765544
step: 30, loss: 0.020679937675595284
step: 40, loss: 0.009657951071858406
step: 50, loss: 0.012889436446130276
step: 60, loss: 0.006811486557126045
step: 70, loss: 0.008199627511203289
step: 80, loss: 0.0032841283828020096
step: 90, loss: 0.08448174595832825
step: 100, loss: 0.05759904906153679
step: 110, loss: 0.08283467590808868
step: 120, loss: 0.000879524159245193
step: 130, loss: 0.014978736639022827
step: 140, loss: 0.00427668821066618
step: 150, loss: 0.030560864135622978
step: 160, loss: 0.09218788146972656
step: 170, loss: 0.00917276181280613
step: 180, loss: 0.004552057478576899
step: 190, loss: 0.0025390007067471743
step: 200, loss: 0.031269509345293045
step: 210, loss: 0.1613999605178833
step: 220, loss: 0.0071815489791333675
step: 230, loss: 0.06160494312644005
step: 240, loss: 0.00987508799880743
step: 250, loss: 0.12924101948738098
step: 260, loss: 0.0026537058874964714
step: 270, loss: 0.0806698426604271
step: 280, loss: 0.11370056867599487
step: 290, loss: 0.03795306012034416
step: 300, loss: 0.023649973794817924
step: 310, loss: 0.08483918756246567
step: 320, loss: 0.0032216007821261883
step: 330, loss: 0.0037346426397562027
step: 340, loss: 0.0070657976903021336
step: 350, loss: 0.0048961034044623375
step: 360, loss: 0.024276966229081154
step: 370, loss: 0.0065468959510326385
step: 380, loss: 0.021091513335704803
```

```

step: 390, loss: 0.029259657487273216
step: 400, loss: 0.027950067073106766
step: 410, loss: 0.005062004551291466
step: 420, loss: 0.020648209378123283
step: 430, loss: 0.10759465396404266
step: 440, loss: 0.036981746554374695
step: 450, loss: 0.059811074286699295
step: 460, loss: 0.01614619791507721
step: 470, loss: 0.0346820205450058
step: 480, loss: 0.008912820369005203
step: 490, loss: 0.007226788438856602
step: 500, loss: 0.03792177513241768
step: 510, loss: 0.016789710149168968
step: 520, loss: 0.07044201344251633
step: 530, loss: 0.025195740163326263
/usr/local/lib/python3.7/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=0.9085872576177285, f1=0.9114391143911439,
best_f1=0.9094304388422036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.18542005121707916
step: 10, loss: 0.06727973371744156
step: 20, loss: 0.0781380906701088
step: 30, loss: 0.0019179201917722821
step: 40, loss: 0.004981784150004387
step: 50, loss: 0.006964416708797216
step: 60, loss: 0.03579100966453552
step: 70, loss: 0.03727424144744873
step: 80, loss: 0.0013326979242265224
step: 90, loss: 0.04618821293115616
step: 100, loss: 0.013198843225836754
step: 110, loss: 0.022630246356129646
step: 120, loss: 0.0032452959567308426
step: 130, loss: 0.000993763329461217
step: 140, loss: 0.004526820965111256
step: 150, loss: 0.0007090538274496794
step: 160, loss: 0.0010285477619618177
step: 170, loss: 0.014635191299021244
step: 180, loss: 0.0003215721226297319
step: 190, loss: 0.0018574486020952463
step: 200, loss: 0.00332138454541564
step: 210, loss: 0.010706671513617039
step: 220, loss: 0.0028332474175840616
step: 230, loss: 0.0044870986603200436
step: 240, loss: 0.00793865229934454

```

```

step: 250, loss: 0.0006601769127883017
step: 260, loss: 0.005956884939223528
step: 270, loss: 0.017000148072838783
step: 280, loss: 0.008502914570271969
step: 290, loss: 0.22260624170303345
step: 300, loss: 0.004804960917681456
step: 310, loss: 0.0014086677692830563
step: 320, loss: 0.004899685271084309
step: 330, loss: 0.015576068311929703
step: 340, loss: 0.0026593529619276524
step: 350, loss: 0.011866766959428787
step: 360, loss: 0.014502612873911858
step: 370, loss: 0.0928075909614563
step: 380, loss: 0.0016682625282555819
step: 390, loss: 0.000428446801379323
step: 400, loss: 0.02271287888288498
step: 410, loss: 0.0008707576198503375
step: 420, loss: 0.002678428776562214
step: 430, loss: 0.001983599504455924
step: 440, loss: 0.014816178008913994
step: 450, loss: 0.0014469235902652144
step: 460, loss: 0.009423049166798592
step: 470, loss: 0.0009958164300769567
step: 480, loss: 0.016396719962358475
step: 490, loss: 0.0006591746932826936
step: 500, loss: 0.012513939291238785
step: 510, loss: 0.045517925173044205
step: 520, loss: 0.004745967220515013
step: 530, loss: 0.0008910807082429528
/usr/local/lib/python3.7/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=0.9138823529411764, f1=0.9066918001885014,
best_f1=0.9094304388422036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0025951277930289507
step: 10, loss: 0.0010137355420738459
step: 20, loss: 0.058443523943424225
step: 30, loss: 0.014291886240243912
step: 40, loss: 0.029803873971104622
step: 50, loss: 0.02585124596953392
step: 60, loss: 0.0013539539650082588
step: 70, loss: 0.006077876314520836
step: 80, loss: 0.0017450586892664433
step: 90, loss: 0.0019030902767553926
step: 100, loss: 0.0034091060515493155

```

```
step: 110, loss: 0.002590242540463805
step: 120, loss: 0.0012333259219303727
step: 130, loss: 0.00129959755577147
step: 140, loss: 0.18697185814380646
step: 150, loss: 0.0011780220083892345
step: 160, loss: 0.005205962341278791
step: 170, loss: 0.0038690040819346905
step: 180, loss: 0.00026012014131993055
step: 190, loss: 0.00363499135710299
step: 200, loss: 0.002284741261973977
step: 210, loss: 0.0024792090989649296
step: 220, loss: 0.002287053968757391
step: 230, loss: 0.0004852268612012267
step: 240, loss: 0.003590776352211833
step: 250, loss: 0.010212046094238758
step: 260, loss: 0.0006417833501473069
step: 270, loss: 0.11922129988670349
step: 280, loss: 0.08108306676149368
step: 290, loss: 0.0009587758686393499
step: 300, loss: 0.0003522929619066417
step: 310, loss: 0.0005858317017555237
step: 320, loss: 0.00046292191836982965
step: 330, loss: 0.0024228328838944435
step: 340, loss: 0.15088613331317902
step: 350, loss: 0.008663907647132874
step: 360, loss: 0.0074203768745064735
step: 370, loss: 0.029101891443133354
step: 380, loss: 0.007404194213449955
step: 390, loss: 0.003193833166733384
step: 400, loss: 0.025502949953079224
step: 410, loss: 0.0015063921455293894
step: 420, loss: 0.011638288386166096
step: 430, loss: 0.0010246913880109787
step: 440, loss: 0.0003254944458603859
step: 450, loss: 0.000548487703781575
step: 460, loss: 0.00017885270062834024
step: 470, loss: 0.008978284895420074
step: 480, loss: 0.002614029450342059
step: 490, loss: 0.0035849595442414284
step: 500, loss: 0.10534856468439102
step: 510, loss: 0.0031874768901616335
step: 520, loss: 0.10527743399143219
step: 530, loss: 0.021672924980521202
/usr/local/lib/python3.7/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=0.9096864763687412, f1=0.9140989729225022,
best_f1=0.9094304388422036
```

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

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

step: 0, loss: 0.0006613505538553
step: 10, loss: 0.0012190265115350485
step: 20, loss: 0.0014491943875327706
step: 30, loss: 0.002621658146381378
step: 40, loss: 0.003953542094677687
step: 50, loss: 0.0022209591697901487
step: 60, loss: 0.07669061422348022
step: 70, loss: 0.037964705377817154
step: 80, loss: 0.0669066458940506
step: 90, loss: 0.0012577433371916413
step: 100, loss: 0.0008769165142439306
step: 110, loss: 0.0005076148081570864
step: 120, loss: 0.0005082474090158939
step: 130, loss: 0.0006627518450841308
step: 140, loss: 0.001256463467143476
step: 150, loss: 0.0026942577678710222
step: 160, loss: 0.00010433196439407766
step: 170, loss: 0.0005914315115660429
step: 180, loss: 0.0009070396772585809
step: 190, loss: 0.0006506310892291367
step: 200, loss: 0.0006909586954861879
step: 210, loss: 0.13477732241153717
step: 220, loss: 0.0002372202288825065
step: 230, loss: 0.17931558191776276
step: 240, loss: 0.00422305054962635
step: 250, loss: 0.0012137752491980791
step: 260, loss: 0.01871895045042038
step: 270, loss: 0.00017950274923350662
step: 280, loss: 0.0021252771839499474
step: 290, loss: 0.004264416638761759
step: 300, loss: 0.0004773074761033058
step: 310, loss: 0.02889949455857277
step: 320, loss: 0.002586348680779338
step: 330, loss: 0.0002107762120431289
step: 340, loss: 0.0018796140793710947
step: 350, loss: 0.002318108920007944
step: 360, loss: 0.0012525967322289944
step: 370, loss: 0.0023347234819084406
step: 380, loss: 0.0008434999617747962
step: 390, loss: 0.00025965546956285834
step: 400, loss: 0.0005829064175486565
step: 410, loss: 0.0059595247730612755
step: 420, loss: 0.004065598826855421
step: 430, loss: 0.006092015188187361
step: 440, loss: 0.00023194565437734127

```

step: 450, loss: 0.00021498952992260456
step: 460, loss: 0.0006896554259583354
step: 470, loss: 0.0025351662188768387
step: 480, loss: 0.2095939964056015
step: 490, loss: 0.0022275305818766356
step: 500, loss: 0.0006884145550429821
step: 510, loss: 0.017464011907577515
step: 520, loss: 0.038664624094963074
step: 530, loss: 0.0012507871724665165
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9206349206349207, f1=0.9217148182665423,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005650388775393367
step: 10, loss: 0.0018671437865123153
step: 20, loss: 0.0040298751555383205
step: 30, loss: 0.0030187792144715786
step: 40, loss: 0.0037945189978927374
step: 50, loss: 0.020234690979123116
step: 60, loss: 0.001059631584212184
step: 70, loss: 0.007919171825051308
step: 80, loss: 0.006039959378540516
step: 90, loss: 0.0010351224336773157
step: 100, loss: 0.0010892859427258372
step: 110, loss: 0.00020653374667745084
step: 120, loss: 0.0008611121447756886
step: 130, loss: 0.00022201014508027583
step: 140, loss: 0.009934407658874989
step: 150, loss: 0.00039981905138120055
step: 160, loss: 0.014933893457055092
step: 170, loss: 0.0985182374715805
step: 180, loss: 0.0038950778543949127
step: 190, loss: 0.00824741180986166
step: 200, loss: 0.0007228046888485551
step: 210, loss: 0.0011049280874431133
step: 220, loss: 0.005849468521773815
step: 230, loss: 0.02762734703719616
step: 240, loss: 0.0007493699085898697
step: 250, loss: 6.621843203902245e-05
step: 260, loss: 0.0008840385708026588

```

```

step: 270, loss: 0.0022149852011352777
step: 280, loss: 0.04242132231593132
step: 290, loss: 9.16691351449117e-05
step: 300, loss: 0.009626622311770916
step: 310, loss: 0.0005585594335570931
step: 320, loss: 0.012573247775435448
step: 330, loss: 0.003999853506684303
step: 340, loss: 0.0004972160677425563
step: 350, loss: 0.00018910481594502926
step: 360, loss: 0.012378744781017303
step: 370, loss: 8.706390508450568e-05
step: 380, loss: 0.0019927467219531536
step: 390, loss: 0.0671495720744133
step: 400, loss: 0.0001457906182622537
step: 410, loss: 0.00033553093089722097
step: 420, loss: 0.0004097643541172147
step: 430, loss: 0.0008826041594147682
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 440, loss: 0.0002603552129585296
step: 450, loss: 0.0006157233146950603
step: 460, loss: 0.00023841277288738638
step: 470, loss: 0.00029017950873821974
step: 480, loss: 0.0009677457856014371
step: 490, loss: 0.00027685289387591183
step: 500, loss: 0.001551009132526815
step: 510, loss: 0.0003515036078169942
step: 520, loss: 0.0008340540225617588
step: 530, loss: 8.716130105312914e-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=0.9166666666666667, f1=0.9168949771689499,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006763663841411471
step: 10, loss: 9.939727169694379e-05
step: 20, loss: 0.0003086072101723403
step: 30, loss: 0.0004994016489945352
step: 40, loss: 7.849602843634784e-05
step: 50, loss: 6.0797890910180286e-05
step: 60, loss: 4.2525469325482845e-05
step: 70, loss: 0.00018514586554374546
step: 80, loss: 0.000295343721518293
step: 90, loss: 9.431268699700013e-05
step: 100, loss: 9.560255421092734e-05
step: 110, loss: 0.00011171709047630429

```



```
step: 120, loss: 6.419309647753835e-05
step: 130, loss: 0.001816938747651875
step: 140, loss: 0.000837742059957236
step: 150, loss: 7.871480192989111e-05
step: 160, loss: 0.0009772501653060317
step: 170, loss: 0.013661985285580158
step: 180, loss: 7.174986239988357e-05
step: 190, loss: 0.0016980392392724752
step: 200, loss: 0.00011108517355751246
step: 210, loss: 8.8558103016112e-05
step: 220, loss: 0.00033863188582472503
step: 230, loss: 6.232516898307949e-05
step: 240, loss: 9.22418330446817e-05
step: 250, loss: 0.0005831416929140687
step: 260, loss: 0.04116189107298851
step: 270, loss: 5.734899968956597e-05
step: 280, loss: 7.135972555261105e-05
step: 290, loss: 0.0005692597478628159
step: 300, loss: 0.0001064229363691993
step: 310, loss: 0.00031102797947824
step: 320, loss: 0.00011119338887510821
step: 330, loss: 0.055714935064315796
step: 340, loss: 7.10840686224401e-05
step: 350, loss: 0.0077296169474720955
step: 360, loss: 5.6981269153766334e-05
step: 370, loss: 0.0007043425575830042
step: 380, loss: 0.0003625699901022017
step: 390, loss: 0.00023487793805543333
step: 400, loss: 0.00041448144474998116
step: 410, loss: 6.568709795828909e-05
step: 420, loss: 0.00010407637455500662
step: 430, loss: 0.003749329363927245
step: 440, loss: 0.002058332087472081
step: 450, loss: 9.323958511231467e-05
step: 460, loss: 0.001286929240450263
step: 470, loss: 3.388798359083012e-05
step: 480, loss: 7.73786596255377e-05
step: 490, loss: 0.00012157404125900939
step: 500, loss: 0.002379857935011387
step: 510, loss: 0.00010700523853302002
step: 520, loss: 0.00016060253256000578
step: 530, loss: 5.2331281040096655e-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=0.912330051570558, f1=0.9187180678123548,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
```

An input tensor was not cuda.

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

step: 0, loss: 8.255136344814673e-05
step: 10, loss: 6.30632130196318e-05
step: 20, loss: 0.00020757026504725218
step: 30, loss: 0.0013222479028627276
step: 40, loss: 0.00010134552576346323
step: 50, loss: 0.000976894749328494
step: 60, loss: 5.438566586235538e-05
step: 70, loss: 5.3982359531801194e-05
step: 80, loss: 7.659979746676981e-05
step: 90, loss: 6.959021266084164e-05
step: 100, loss: 7.551562885055318e-05
step: 110, loss: 4.9069232773035765e-05
step: 120, loss: 0.0003717165964189917
step: 130, loss: 0.00010920553904725239
step: 140, loss: 0.0015061793383210897
step: 150, loss: 0.00027083721943199635
step: 160, loss: 6.40108555671759e-05
step: 170, loss: 0.00011403641838114709
step: 180, loss: 7.215445657493547e-05
step: 190, loss: 8.870491728885099e-05
step: 200, loss: 0.0004221806884743273
step: 210, loss: 0.0011637161951512098
step: 220, loss: 0.01908283866941929
step: 230, loss: 0.00020676500571426004
step: 240, loss: 0.00016398528532590717
step: 250, loss: 5.245941429166123e-05
step: 260, loss: 0.0006227224366739392
step: 270, loss: 0.00016089082055259496
step: 280, loss: 0.0004532256571110338
step: 290, loss: 0.00011889773304574192
step: 300, loss: 4.1318849980598316e-05
step: 310, loss: 0.038059089332818985
step: 320, loss: 0.0002890423056669533
step: 330, loss: 4.164514393778518e-05
step: 340, loss: 0.00036558497231453657
step: 350, loss: 0.00025720111443661153
step: 360, loss: 0.0001996352366404608
step: 370, loss: 0.0029504194390028715
step: 380, loss: 0.00013740982103627175
step: 390, loss: 0.0012242940720170736
step: 400, loss: 0.00012197878095321357
step: 410, loss: 0.000264778733253479
step: 420, loss: 0.0007940392242744565
step: 430, loss: 3.589110201573931e-05
step: 440, loss: 0.011596345342695713
step: 450, loss: 5.3007519454695284e-05

```

step: 460, loss: 3.9887996535981074e-05
step: 470, loss: 9.435562242288142e-05
step: 480, loss: 0.00011534568329807371
step: 490, loss: 0.0005237425793893635
step: 500, loss: 0.006045225076377392
step: 510, loss: 3.7392597732832655e-05
step: 520, loss: 8.660013554617763e-05
step: 530, loss: 0.00015465117758139968
/usr/local/lib/python3.7/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=0.9109004739336494, f1=0.9116541353383459,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00040583443478681147
step: 10, loss: 0.0017220880836248398
step: 20, loss: 7.135054329410195e-05
step: 30, loss: 0.0007998099317774177
step: 40, loss: 0.03720865026116371
step: 50, loss: 5.8461810112930834e-05
step: 60, loss: 0.0008020593668334186
step: 70, loss: 0.00012480041186790913
step: 80, loss: 4.009261101600714e-05
step: 90, loss: 9.656733163865283e-05
step: 100, loss: 0.002373303985223174
step: 110, loss: 0.00025161131634376943
step: 120, loss: 4.2830462916754186e-05
step: 130, loss: 4.398117016535252e-05
step: 140, loss: 0.00033955954131670296
step: 150, loss: 2.621045678097289e-05
step: 160, loss: 2.9626427931361832e-05
step: 170, loss: 5.567954940488562e-05
step: 180, loss: 5.974435043754056e-05
step: 190, loss: 0.0003876701812259853
step: 200, loss: 0.07276774942874908
step: 210, loss: 4.6666253183502704e-05
step: 220, loss: 0.01226213201880455
step: 230, loss: 0.01772150956094265
step: 240, loss: 0.0001030282219289802
step: 250, loss: 5.022447294322774e-05
step: 260, loss: 3.846890467684716e-05
step: 270, loss: 9.751049219630659e-05
step: 280, loss: 0.0001469978888053447
step: 290, loss: 5.8799239923246205e-05
step: 300, loss: 0.0012087946524843574
step: 310, loss: 0.00046300876419991255

```

```

step: 320, loss: 0.00037817409611307085
step: 330, loss: 7.045833626762033e-05
step: 340, loss: 0.0005904326098971069
step: 350, loss: 0.00023739435710012913
step: 360, loss: 2.607981878099963e-05
step: 370, loss: 0.0013802580069750547
step: 380, loss: 4.665645974455401e-05
step: 390, loss: 4.35433721577283e-05
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 7.441703928634524e-05
step: 410, loss: 8.693610288901255e-05
step: 420, loss: 0.00031432227115146816
step: 430, loss: 5.702791895600967e-05
step: 440, loss: 7.774481491651386e-05
step: 450, loss: 6.3696141296532e-05
step: 460, loss: 0.05062633752822876
step: 470, loss: 0.0001053093874361366
step: 480, loss: 0.00010337492130929604
step: 490, loss: 6.769036554032937e-05
step: 500, loss: 0.008400246500968933
step: 510, loss: 4.694828385254368e-05
step: 520, loss: 0.0001598810777068138
step: 530, loss: 0.00016703479923307896
/usr/local/lib/python3.7/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=0.9172093023255814, f1=0.9126213592233011,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.121121310163289e-05
step: 10, loss: 0.00010874636063817888
step: 20, loss: 4.8470988986082375e-05
step: 30, loss: 4.488399645197205e-05
step: 40, loss: 4.43886783614289e-05
step: 50, loss: 0.00573759013786912
step: 60, loss: 0.0012242953525856137
step: 70, loss: 0.00012600512127391994
step: 80, loss: 2.7972724637947977e-05
step: 90, loss: 8.51728837005794e-05
step: 100, loss: 0.00015457489644177258
step: 110, loss: 4.18292147514876e-05
step: 120, loss: 3.0602532206103206e-05
step: 130, loss: 0.00010600533278193325
step: 140, loss: 3.6189983802614734e-05
step: 150, loss: 3.1496540032094344e-05
step: 160, loss: 3.600010313675739e-05

```

```

step: 170, loss: 0.00018579259631223977
step: 180, loss: 4.14560163335409e-05
step: 190, loss: 8.503046410623938e-05
step: 200, loss: 4.631328920368105e-05
step: 210, loss: 6.216738256625831e-05
step: 220, loss: 8.722685015527532e-05
step: 230, loss: 0.0001610312465345487
step: 240, loss: 0.00028941838536411524
step: 250, loss: 5.167450581211597e-05
step: 260, loss: 3.1202078389469534e-05
step: 270, loss: 5.01617141708266e-05
step: 280, loss: 6.109073729021475e-05
step: 290, loss: 3.514694617479108e-05
step: 300, loss: 0.0001582438126206398
step: 310, loss: 5.987327313050628e-05
step: 320, loss: 3.086296055698767e-05
step: 330, loss: 0.00013075278548058122
step: 340, loss: 2.962644248327706e-05
step: 350, loss: 8.727169188205153e-05
step: 360, loss: 0.00011677419388433918
step: 370, loss: 4.6538960305042565e-05
step: 380, loss: 3.241653030272573e-05
step: 390, loss: 3.024119359906763e-05
step: 400, loss: 0.00012174983567092568
step: 410, loss: 0.0004303231544326991
step: 420, loss: 0.000520272005815059
step: 430, loss: 0.000365823827451095
step: 440, loss: 3.713175465236418e-05
step: 450, loss: 4.4581662223208696e-05
step: 460, loss: 2.7033685910282657e-05
step: 470, loss: 8.082083513727412e-05
step: 480, loss: 0.0017260446911677718
step: 490, loss: 0.00043435022234916687
step: 500, loss: 3.873186869896017e-05
step: 510, loss: 8.6315514636226e-05
step: 520, loss: 3.0691571737406775e-05
step: 530, loss: 4.2420633690198883e-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=0.9111424541607899, f1=0.9119025304592315,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021738240029662848
step: 10, loss: 0.004684855230152607
step: 20, loss: 6.286859570536762e-05

```

step: 30, loss: 6.389882764779031e-05
step: 40, loss: 2.133071029675193e-05
step: 50, loss: 3.5813147405860946e-05
step: 60, loss: 3.320858741062693e-05
step: 70, loss: 3.6467696190811694e-05
step: 80, loss: 1.7579348423168994e-05
step: 90, loss: 0.0004220016417093575
step: 100, loss: 7.744241156615317e-05
step: 110, loss: 2.2246800654102117e-05
step: 120, loss: 2.702602796489373e-05
step: 130, loss: 2.07493740163045e-05
step: 140, loss: 4.394137067720294e-05
step: 150, loss: 0.07993196696043015
step: 160, loss: 2.032472548307851e-05
step: 170, loss: 0.00018914291285909712
step: 180, loss: 4.059444472659379e-05
step: 190, loss: 7.513039599871263e-05
step: 200, loss: 3.2874333555810153e-05
step: 210, loss: 4.548911238089204e-05
step: 220, loss: 2.8504793590400368e-05
step: 230, loss: 1.9173761756974272e-05
step: 240, loss: 3.786956949625164e-05
step: 250, loss: 2.3044171030051075e-05
step: 260, loss: 1.9158876966685057e-05
step: 270, loss: 2.0473746189964004e-05
step: 280, loss: 2.7953568860539235e-05
step: 290, loss: 1.5839763364056125e-05
step: 300, loss: 2.640049933688715e-05
step: 310, loss: 0.001467360183596611
step: 320, loss: 0.0015486617339774966
step: 330, loss: 0.007711879443377256
step: 340, loss: 2.8895448849652894e-05
step: 350, loss: 2.2366204575519077e-05
step: 360, loss: 3.4864991903305054e-05
step: 370, loss: 5.1317663746885955e-05
step: 380, loss: 0.0012646878603845835
step: 390, loss: 0.0003970117249991745
step: 400, loss: 2.4370394385186955e-05
step: 410, loss: 5.979060006211512e-05
step: 420, loss: 4.6525474317604676e-05
step: 430, loss: 0.051170893013477325
step: 440, loss: 0.0001008570397971198
step: 450, loss: 9.679418144514784e-05
step: 460, loss: 0.0023260649759322405
step: 470, loss: 5.4264775826595724e-05
step: 480, loss: 2.9108448870829307e-05
step: 490, loss: 2.8441088943509385e-05
step: 500, loss: 2.0622153897420503e-05

```

step: 510, loss: 2.1714222384616733e-05
step: 520, loss: 0.0045199403539299965
step: 530, loss: 0.0001350540405837819
/usr/local/lib/python3.7/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=0.9110251450676983, f1=0.901884968583857,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.86422349140048e-05
step: 10, loss: 2.0440129446797073e-05
step: 20, loss: 2.3315413272939622e-05
step: 30, loss: 2.0894798581139185e-05
step: 40, loss: 3.069510421482846e-05
step: 50, loss: 0.00010482859215699136
step: 60, loss: 1.6703921573935077e-05
step: 70, loss: 0.0012391365598887205
step: 80, loss: 3.514172931318171e-05
step: 90, loss: 2.3796676032361574e-05
step: 100, loss: 3.457610364421271e-05
step: 110, loss: 1.9937464458053e-05
step: 120, loss: 3.917906724382192e-05
step: 130, loss: 0.00032343470957130194
step: 140, loss: 1.9043387510464527e-05
step: 150, loss: 1.4323602044896688e-05
step: 160, loss: 2.0090117686777376e-05
step: 170, loss: 3.4016542485915124e-05
step: 180, loss: 2.2824089683126658e-05
step: 190, loss: 5.14735329488758e-05
step: 200, loss: 2.5256862500100397e-05
step: 210, loss: 2.0186940673738718e-05
step: 220, loss: 0.00016572764434386045
step: 230, loss: 3.8355017750291154e-05
step: 240, loss: 3.229077628930099e-05
step: 250, loss: 2.28057979256846e-05
step: 260, loss: 8.492390770697966e-05
step: 270, loss: 3.0205857910914347e-05
step: 280, loss: 1.9803253962891176e-05
step: 290, loss: 0.00018728923168964684
step: 300, loss: 2.8620097509701736e-05
step: 310, loss: 0.0013218637323006988
step: 320, loss: 4.0396236727247015e-05
step: 330, loss: 0.00024313764879480004
step: 340, loss: 3.277305222582072e-05
step: 350, loss: 4.1815870645223185e-05
step: 360, loss: 0.0003849854110740125

```

```
step: 370, loss: 3.5779961763182655e-05
step: 380, loss: 0.00011059325333917513
step: 390, loss: 0.011282576248049736
step: 400, loss: 0.0011931052431464195
step: 410, loss: 2.9335238650674e-05
step: 420, loss: 5.193801553105004e-05
step: 430, loss: 2.9175558665883727e-05
step: 440, loss: 0.00028444171766750515
step: 450, loss: 0.0014502472477033734
step: 460, loss: 2.1662192011717707e-05
step: 470, loss: 3.656798435258679e-05
step: 480, loss: 3.734367055585608e-05
step: 490, loss: 8.57944687595591e-05
step: 500, loss: 1.6759891877882183e-05
step: 510, loss: 0.008315177634358406
step: 520, loss: 3.807796383625828e-05
step: 530, loss: 0.00010337359708501026
/usr/local/lib/python3.7/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=0.9146853146853147, f1=0.9185185185185185,
best_f1=0.9217148182665423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.6955602152156644e-05
step: 10, loss: 2.8318325348664075e-05
step: 20, loss: 3.801352795562707e-05
step: 30, loss: 0.00018729858857113868
step: 40, loss: 0.047205738723278046
step: 50, loss: 9.551606490276754e-05
step: 60, loss: 2.5089406335609965e-05
step: 70, loss: 0.00027408695314079523
step: 80, loss: 9.521297761239111e-05
step: 90, loss: 2.845646122295875e-05
step: 100, loss: 4.0972401620820165e-05
step: 110, loss: 9.221676737070084e-05
step: 120, loss: 0.0010312494123354554
step: 130, loss: 0.0049126967787742615
step: 140, loss: 2.3200573195936158e-05
step: 150, loss: 2.808809767884668e-05
step: 160, loss: 2.9525750505854376e-05
step: 170, loss: 2.496572778909467e-05
step: 180, loss: 1.867462196969427e-05
step: 190, loss: 6.292795296758413e-05
step: 200, loss: 3.0169592719175853e-05
step: 210, loss: 0.0004147505678702146
step: 220, loss: 2.498106005077716e-05
```



```

step: 230, loss: 2.589371615613345e-05
step: 240, loss: 2.7785548809333704e-05
step: 250, loss: 2.4586039216956124e-05
step: 260, loss: 1.985923881875351e-05
step: 270, loss: 2.108480293827597e-05
step: 280, loss: 7.471974095096812e-05
step: 290, loss: 1.8749149603536353e-05
step: 300, loss: 2.3584174414281733e-05
step: 310, loss: 4.1689942008815706e-05
step: 320, loss: 0.00016299469280056655
step: 330, loss: 0.0002987776242662221
step: 340, loss: 2.8369922802085057e-05
step: 350, loss: 2.6321706172893755e-05
step: 360, loss: 0.00017333502182736993
step: 370, loss: 1.9006141883437522e-05
step: 380, loss: 5.792810770799406e-05
step: 390, loss: 9.187281830236316e-05
step: 400, loss: 3.503073821775615e-05
step: 410, loss: 2.550215504015796e-05
step: 420, loss: 0.007548716850578785
step: 430, loss: 8.820556104183197e-05
step: 440, loss: 3.1961561035132036e-05
step: 450, loss: 2.1509449652512558e-05
step: 460, loss: 2.488409518264234e-05
step: 470, loss: 0.008780020289123058
step: 480, loss: 1.5277260899893008e-05
step: 490, loss: 1.8741669919108972e-05
step: 500, loss: 2.7525205950951204e-05
step: 510, loss: 3.514491982059553e-05
step: 520, loss: 3.206974361091852e-05
step: 530, loss: 3.35348850057926e-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=0.9123134328358209, f1=0.9181692094313454,
best_f1=0.9217148182665423
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, 269.46it/s]
load_f1 = 0.9218604651162791
real_f1 = 0.9205020920502092
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, 270.88it/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 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: 0.5511749982833862

```

step: 10, loss: 0.366182416677475
step: 20, loss: 0.3707675039768219
step: 30, loss: 0.33046072721481323
step: 40, loss: 0.2013465315103531
step: 50, loss: 0.44803348183631897
step: 60, loss: 0.33485835790634155
step: 70, loss: 0.23659294843673706
step: 80, loss: 0.20745640993118286
step: 90, loss: 0.46275436878204346
step: 100, loss: 0.5162782669067383
step: 110, loss: 0.2588721513748169
step: 120, loss: 0.21401843428611755
step: 130, loss: 0.26725494861602783
step: 140, loss: 0.25542283058166504
step: 150, loss: 0.3431297540664673
step: 160, loss: 0.29818177223205566
step: 170, loss: 0.31328216195106506
step: 180, loss: 0.1020389273762703
step: 190, loss: 0.19218020141124725
step: 200, loss: 0.3173246383666992
step: 210, loss: 0.2077813446521759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.5141843971631205, f1=0.5549738219895288,
best_f1=0.5549738219895288
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.13466346263885498
step: 10, loss: 0.2412434071302414
step: 20, loss: 0.17177319526672363
step: 30, loss: 0.19241701066493988
step: 40, loss: 0.17342735826969147
step: 50, loss: 0.1817523092031479
step: 60, loss: 0.416681706905365
step: 70, loss: 0.1303916871547699
step: 80, loss: 0.24903303384780884
step: 90, loss: 0.14727701246738434
step: 100, loss: 0.060544826090335846
step: 110, loss: 0.11061209440231323
step: 120, loss: 0.1754954308271408
step: 130, loss: 0.0455249659717083
step: 140, loss: 0.21849335730075836

```

```

step: 150, loss: 0.28749752044677734
step: 160, loss: 0.19872166216373444
step: 170, loss: 0.17957235872745514
step: 180, loss: 0.22695370018482208
step: 190, loss: 0.1942475438117981
step: 200, loss: 0.06629502773284912
step: 210, loss: 0.09032202512025833
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5784753363228701, f1=0.501240694789082,
best_f1=0.501240694789082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06278422474861145
step: 10, loss: 0.20109079778194427
step: 20, loss: 0.16583503782749176
step: 30, loss: 0.2854541540145874
step: 40, loss: 0.046691231429576874
step: 50, loss: 0.12370912730693817
step: 60, loss: 0.2065567523241043
step: 70, loss: 0.20588736236095428
step: 80, loss: 0.11195869743824005
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 90, loss: 0.05866504833102226
step: 100, loss: 0.10154219716787338
step: 110, loss: 0.14239725470542908
step: 120, loss: 0.22616414725780487
step: 130, loss: 0.15047594904899597
step: 140, loss: 0.08509472757577896
step: 150, loss: 0.2119530439376831
step: 160, loss: 0.05866415798664093
step: 170, loss: 0.08840569853782654
step: 180, loss: 0.058072708547115326
step: 190, loss: 0.16707001626491547
step: 200, loss: 0.045936521142721176
step: 210, loss: 0.0973811224102974
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5848375451263538, f1=0.596745027124774,
best_f1=0.596745027124774
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.320909321308136
step: 10, loss: 0.12331082671880722
step: 20, loss: 0.10462630540132523
step: 30, loss: 0.3499193489551544
step: 40, loss: 0.06854365020990372
step: 50, loss: 0.15896283090114594
step: 60, loss: 0.14231239259243011
step: 70, loss: 0.3141816258430481
step: 80, loss: 0.13184918463230133
step: 90, loss: 0.05803495645523071
step: 100, loss: 0.27958810329437256
step: 110, loss: 0.09275826811790466
step: 120, loss: 0.13413439691066742
step: 130, loss: 0.1624361276626587
step: 140, loss: 0.23726606369018555
step: 150, loss: 0.08997891843318939
step: 160, loss: 0.033671360462903976
step: 170, loss: 0.08825062960386276
step: 180, loss: 0.33008164167404175
step: 190, loss: 0.05936490371823311
step: 200, loss: 0.18463142216205597
step: 210, loss: 0.1599002629518509
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5996055226824457, f1=0.5806451612903225,
best_f1=0.5806451612903225
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.1953476220369339
step: 10, loss: 0.025740928947925568
step: 20, loss: 0.21210703253746033
step: 30, loss: 0.033502157777547836
step: 40, loss: 0.04560543969273567
step: 50, loss: 0.12349800020456314
step: 60, loss: 0.156165212392807
step: 70, loss: 0.26554664969444275

```

```

step: 80, loss: 0.040700070559978485
step: 90, loss: 0.05584985762834549
step: 100, loss: 0.010340450331568718
step: 110, loss: 0.15664911270141602
step: 120, loss: 0.1538524329662323
step: 130, loss: 0.12406452745199203
step: 140, loss: 0.05935625359416008
step: 150, loss: 0.040500398725271225
step: 160, loss: 0.14283747971057892
step: 170, loss: 0.06873826682567596
step: 180, loss: 0.07027847319841385
step: 190, loss: 0.039688531309366226
step: 200, loss: 0.0688035860657692
step: 210, loss: 0.038344766944646835
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.610752688172043, f1=0.5720430107526882,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03535377234220505
step: 10, loss: 0.07975593209266663
step: 20, loss: 0.07366737723350525
step: 30, loss: 0.004489730577915907
step: 40, loss: 0.08611752092838287
step: 50, loss: 0.011642887257039547
step: 60, loss: 0.08366288244724274
step: 70, loss: 0.0217424426227808
step: 80, loss: 0.019633665680885315
step: 90, loss: 0.07630892097949982
step: 100, loss: 0.007801597472280264
step: 110, loss: 0.0033159914892166853
step: 120, loss: 0.021465396508574486
step: 130, loss: 0.008629241026937962
step: 140, loss: 0.13476857542991638
step: 150, loss: 0.03208759427070618
step: 160, loss: 0.02214445173740387
step: 170, loss: 0.14349918067455292
step: 180, loss: 0.06041873246431351
step: 190, loss: 0.10130239278078079
step: 200, loss: 0.00831328984349966
step: 210, loss: 0.01715260185301304

```

```

/usr/local/lib/python3.7/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=0.5601750547045952, f1=0.5387931034482758,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001117223990149796
step: 10, loss: 0.0026957800146192312
step: 20, loss: 0.00589274475350976
step: 30, loss: 0.2844163179397583
step: 40, loss: 0.009759011678397655
step: 50, loss: 0.07826033979654312
step: 60, loss: 0.01342763565480709
step: 70, loss: 0.037232864648103714
step: 80, loss: 0.03862684220075607
step: 90, loss: 0.15690074861049652
step: 100, loss: 0.0042364210821688175
step: 110, loss: 0.04560106247663498
step: 120, loss: 0.1241108775138855
step: 130, loss: 0.07751170545816422
step: 140, loss: 0.014990096911787987
step: 150, loss: 0.07187534868717194
step: 160, loss: 0.0208045095205307
step: 170, loss: 0.01507838536053896
step: 180, loss: 0.005490211304277182
step: 190, loss: 0.09241436421871185
step: 200, loss: 0.008832084946334362
step: 210, loss: 0.10592106729745865
/usr/local/lib/python3.7/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=0.5189873417721518, f1=0.5324947589098533,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0526016466319561
step: 10, loss: 0.04585672914981842
step: 20, loss: 0.02368415892124176
step: 30, loss: 0.018340809270739555
step: 40, loss: 0.04744570329785347
step: 50, loss: 0.060997433960437775
step: 60, loss: 0.02151540294289589
step: 70, loss: 0.1312265396118164
step: 80, loss: 0.08739916235208511
step: 90, loss: 0.002228862838819623

```



```

step: 100, loss: 0.022967655211687088
step: 110, loss: 0.040423933416604996
step: 120, loss: 0.01045741606503725
step: 130, loss: 0.011757639236748219
step: 140, loss: 0.051656901836395264
step: 150, loss: 0.011319179087877274
step: 160, loss: 0.04932812601327896
step: 170, loss: 0.019752727821469307
step: 180, loss: 0.07047766447067261
step: 190, loss: 0.14845529198646545
step: 200, loss: 0.00043714611092582345
step: 210, loss: 0.022275593131780624
/usr/local/lib/python3.7/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=0.5406593406593406, f1=0.5353982300884956,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.015405288897454739
step: 10, loss: 0.025812968611717224
step: 20, loss: 0.001614188775420189
step: 30, loss: 0.02681724540889263
step: 40, loss: 0.03517087921500206
step: 50, loss: 0.012913281098008156
step: 60, loss: 0.10767042636871338
step: 70, loss: 0.05549526959657669
step: 80, loss: 0.000619060592725873
step: 90, loss: 0.0031624312978237867
step: 100, loss: 0.07968215644359589
step: 110, loss: 0.01930244266986847
step: 120, loss: 0.0066214934922754765
step: 130, loss: 0.0015774712665006518
step: 140, loss: 0.0938350260257721
step: 150, loss: 0.2110792100429535
step: 160, loss: 0.03408084437251091
step: 170, loss: 0.002781183458864689
step: 180, loss: 0.0527152419090271
step: 190, loss: 0.0010315484832972288
step: 200, loss: 0.016842355951666832
step: 210, loss: 0.00094818533398211
/usr/local/lib/python3.7/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=0.5684647302904564, f1=0.5308641975308642,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.026407772675156593
step: 10, loss: 0.050160616636276245
step: 20, loss: 0.0014759297482669353
step: 30, loss: 0.06600965559482574
step: 40, loss: 0.003663807874545455
step: 50, loss: 0.0008399923681281507
step: 60, loss: 0.03282426670193672
step: 70, loss: 0.04336530715227127
step: 80, loss: 0.006788403727114201
step: 90, loss: 0.008040199056267738
step: 100, loss: 0.07859454303979874
step: 110, loss: 0.002083119936287403
step: 120, loss: 0.0005342165823094547
step: 130, loss: 0.1114068254828453
step: 140, loss: 0.0005542529397644103
step: 150, loss: 0.0050441669300198555
step: 160, loss: 0.004208774771541357
step: 170, loss: 0.10038872808218002
step: 180, loss: 0.022134464234113693
step: 190, loss: 0.02448248490691185
step: 200, loss: 0.0008513598586432636
step: 210, loss: 0.054069049656391144
/usr/local/lib/python3.7/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=0.5418502202643173, f1=0.49557522123893805,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23760654032230377
step: 10, loss: 0.019122416153550148
step: 20, loss: 0.0023415363393723965
step: 30, loss: 0.0032497846987098455
step: 40, loss: 0.0011866550194099545
step: 50, loss: 0.0005645027267746627
step: 60, loss: 0.034654758870601654
step: 70, loss: 0.0023983579594641924
step: 80, loss: 0.0061037577688694
step: 90, loss: 0.000979534350335598
step: 100, loss: 0.0003939489251933992
step: 110, loss: 0.024915484711527824
step: 120, loss: 0.017129207029938698
step: 130, loss: 0.010307217948138714
step: 140, loss: 0.0011497247032821178
step: 150, loss: 0.0002875995996873826

```

```

step: 160, loss: 0.0005681785987690091
step: 170, loss: 0.0016369407530874014
step: 180, loss: 0.003328090999275446
step: 190, loss: 0.0010164041304960847
step: 200, loss: 0.017027515918016434
step: 210, loss: 0.00020922160183545202
/usr/local/lib/python3.7/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=0.5520361990950227, f1=0.5059101654846335,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012368780560791492
step: 10, loss: 0.0005627256468869746
step: 20, loss: 0.001323063625022769
step: 30, loss: 0.0032912783790379763
step: 40, loss: 0.01355877984315157
step: 50, loss: 0.0005958823603577912
step: 60, loss: 0.001245562918484211
step: 70, loss: 0.0012533963890746236
step: 80, loss: 0.0002436558424960822
step: 90, loss: 0.00030270760180428624
step: 100, loss: 0.0007891844725236297
step: 110, loss: 0.022283043712377548
step: 120, loss: 0.0006510847597382963
step: 130, loss: 0.0015017502009868622
step: 140, loss: 0.0027388210874050856
step: 150, loss: 0.0003744781424757093
step: 160, loss: 0.0009179514017887414
step: 170, loss: 0.0005963305593468249
step: 180, loss: 0.0007726870244368911
step: 190, loss: 0.007493926212191582
step: 200, loss: 0.0005689905374310911
step: 210, loss: 0.004640186205506325
/usr/local/lib/python3.7/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=0.5622317596566524, f1=0.5539112050739957,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03061828203499317
step: 10, loss: 0.000249088421696797
step: 20, loss: 0.018851127475500107
step: 30, loss: 0.002319689141586423

```

```

step: 40, loss: 0.0010493253357708454
step: 50, loss: 0.007389931008219719
step: 60, loss: 0.0021089590154588223
step: 70, loss: 0.09759610146284103
step: 80, loss: 0.030239837244153023
step: 90, loss: 0.00029463027021847665
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0001255377719644457
step: 110, loss: 0.0010422088671475649
step: 120, loss: 0.00018901104340329766
step: 130, loss: 0.00014375218597706407
step: 140, loss: 0.0006792726344428957
step: 150, loss: 0.008908933028578758
step: 160, loss: 0.003459748812019825
step: 170, loss: 0.001042556599713862
step: 180, loss: 0.019772443920373917
step: 190, loss: 0.0003694502520374954
step: 200, loss: 0.03375299647450447
step: 210, loss: 0.0017098491080105305
/usr/local/lib/python3.7/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=0.5511111111111111, f1=0.5312499999999999,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002203792566433549
step: 10, loss: 0.01918868161737919
step: 20, loss: 0.0027525753248482943
step: 30, loss: 0.003555507166311145
step: 40, loss: 0.0008162501617334783
step: 50, loss: 0.0014963659923523664
step: 60, loss: 0.0001023632867145352
step: 70, loss: 0.001372376922518015
step: 80, loss: 0.000451058178441599
step: 90, loss: 0.0008110243361443281
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 100, loss: 0.00023023574613034725
step: 110, loss: 0.00021613638091366738
step: 120, loss: 0.08912710100412369
step: 130, loss: 0.0011021653190255165
step: 140, loss: 0.0021936462726444006
step: 150, loss: 0.00022204082051757723
step: 160, loss: 0.0005207753274589777
step: 170, loss: 0.000273180368822068
step: 180, loss: 0.060152553021907806
step: 190, loss: 0.0005029558669775724

```

```

step: 200, loss: 0.00016129126015584916
step: 210, loss: 0.003945199307054281
/usr/local/lib/python3.7/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=0.5475113122171946, f1=0.5079365079365079,
best_f1=0.5720430107526882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002087611355818808
step: 10, loss: 0.00013374547415878624
step: 20, loss: 0.0004414421273395419
step: 30, loss: 0.0009411587379872799
step: 40, loss: 0.0003972258709836751
step: 50, loss: 0.00022729436750523746
step: 60, loss: 0.0006089657545089722
step: 70, loss: 0.0006554850842803717
step: 80, loss: 0.0007573581533506513
step: 90, loss: 0.000969260057900101
step: 100, loss: 0.00022452045232057571
step: 110, loss: 0.00016333071107510477
step: 120, loss: 0.07452871650457382
step: 130, loss: 0.0003526149957906455
step: 140, loss: 0.00013668743486050516
step: 150, loss: 0.00022308275219984353
step: 160, loss: 0.0004102396487724036
step: 170, loss: 0.0002238692541141063
step: 180, loss: 0.0020769331604242325
step: 190, loss: 0.000748312333598733
step: 200, loss: 0.0004767839564010501
step: 210, loss: 0.0027972201351076365
/usr/local/lib/python3.7/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=0.5515695067264574, f1=0.5135135135135135,
best_f1=0.5720430107526882
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02: FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2293it [00:04, 460.10it/s]
load_f1 = 0.605543710021322
real_f1 = 0.610752688172043
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, 277.04it/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: 0.551241934299469
step: 10, loss: 0.3442268967628479
step: 20, loss: 0.2754955589771271
step: 30, loss: 0.42331406474113464
step: 40, loss: 0.4329490661621094
step: 50, loss: 0.3074737787246704
step: 60, loss: 0.32702839374542236
step: 70, loss: 0.28265485167503357
step: 80, loss: 0.21767285466194153
step: 90, loss: 0.2704833149909973
step: 100, loss: 0.3323630094528198
step: 110, loss: 0.4107508659362793
step: 120, loss: 0.21550558507442474
step: 130, loss: 0.2025929093360901
step: 140, loss: 0.09187358617782593
step: 150, loss: 0.20400075614452362
step: 160, loss: 0.21865563094615936
step: 170, loss: 0.21478836238384247
step: 180, loss: 0.05956636369228363
step: 190, loss: 0.29916244745254517
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.603448275862069, f1=0.6049046321525886,
best_f1=0.6049046321525886
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.26918888092041016
step: 10, loss: 0.11929988861083984
step: 20, loss: 0.13446810841560364
step: 30, loss: 0.1553235650062561
step: 40, loss: 0.22984234988689423
step: 50, loss: 0.21607540547847748
step: 60, loss: 0.45593971014022827
step: 70, loss: 0.1992189586162567
step: 80, loss: 0.1565658301115036
step: 90, loss: 0.1664285659790039
step: 100, loss: 0.03060387447476387
step: 110, loss: 0.12745165824890137
step: 120, loss: 0.3647100627422333
step: 130, loss: 0.22301313281059265
step: 140, loss: 0.1838911920785904
step: 150, loss: 0.17188382148742676

```



```

step: 160, loss: 0.0469038300216198
step: 170, loss: 0.15868493914604187
step: 180, loss: 0.1781843602657318
step: 190, loss: 0.09572729468345642
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7853403141361257, f1=0.7475247524752476,
best_f1=0.7475247524752476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.040263425558805466
step: 10, loss: 0.41498517990112305
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.10758513957262039
step: 30, loss: 0.0921429693698883
step: 40, loss: 0.08793646842241287
step: 50, loss: 0.28443047404289246
step: 60, loss: 0.03144243359565735
step: 70, loss: 0.19332902133464813
step: 80, loss: 0.10272983461618423
step: 90, loss: 0.12993299961090088
step: 100, loss: 0.010468574240803719
step: 110, loss: 0.014572306536138058
step: 120, loss: 0.08075951039791107
step: 130, loss: 0.020747289061546326
step: 140, loss: 0.02424142323434353
step: 150, loss: 0.17717473208904266
step: 160, loss: 0.048288438469171524
step: 170, loss: 0.0870194062590599
step: 180, loss: 0.04472934454679489
step: 190, loss: 0.09943251311779022
/usr/local/lib/python3.7/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=0.784, f1=0.7923497267759563, best_f1=0.7475247524752476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012507784180343151
step: 10, loss: 0.1455596685409546
step: 20, loss: 0.02687423676252365
step: 30, loss: 0.03559611737728119

```

```

step: 40, loss: 0.04662146791815758
step: 50, loss: 0.027737567201256752
step: 60, loss: 0.015153501182794571
step: 70, loss: 0.12479528784751892
step: 80, loss: 0.25576332211494446
step: 90, loss: 0.017922915518283844
step: 100, loss: 0.0168040469288826
step: 110, loss: 0.005659319460391998
step: 120, loss: 0.13925430178642273
step: 130, loss: 0.13123562932014465
step: 140, loss: 0.05354843661189079
step: 150, loss: 0.11001986265182495
step: 160, loss: 0.025357360020279884
step: 170, loss: 0.21307343244552612
step: 180, loss: 0.02536970004439354
step: 190, loss: 0.16815361380577087
/usr/local/lib/python3.7/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=0.7602905569007263, f1=0.762589928057554,
best_f1=0.7475247524752476
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07823115587234497
step: 10, loss: 0.022330613806843758
step: 20, loss: 0.10472246259450912
step: 30, loss: 0.008697472512722015
step: 40, loss: 0.15376994013786316
step: 50, loss: 0.15108168125152588
step: 60, loss: 0.11902771890163422
step: 70, loss: 0.03295550495386124
step: 80, loss: 0.007712875492870808
step: 90, loss: 0.022803623229265213
step: 100, loss: 0.022391576319932938
step: 110, loss: 0.005040720105171204
step: 120, loss: 0.00699401693418622
step: 130, loss: 0.10202246159315109
step: 140, loss: 0.020743347704410553
step: 150, loss: 0.07879636436700821
step: 160, loss: 0.14121828973293304
step: 170, loss: 0.003691786900162697
step: 180, loss: 0.18582087755203247
step: 190, loss: 0.052441082894802094
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.7887323943661971, f1=0.8057142857142858,
best_f1=0.8057142857142858
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.020304309204220772
step: 10, loss: 0.07706576585769653
step: 20, loss: 0.003451837692409754
step: 30, loss: 0.007443738169968128
step: 40, loss: 0.06954097747802734
step: 50, loss: 0.11632711440324783
step: 60, loss: 0.007442977745085955
step: 70, loss: 0.004386697895824909
step: 80, loss: 0.20977771282196045
step: 90, loss: 0.053717635571956635
step: 100, loss: 0.002976436633616686
step: 110, loss: 0.09932296723127365
step: 120, loss: 0.29791128635406494
step: 130, loss: 0.031226173043251038
step: 140, loss: 0.044157758355140686
step: 150, loss: 0.06911744922399521
step: 160, loss: 0.019175520166754723
step: 170, loss: 0.1988099366426468
step: 180, loss: 0.02370123378932476
step: 190, loss: 0.018042566254734993
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7988826815642459, f1=0.8067226890756303,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0040992246940732
step: 10, loss: 0.02919011190533638
step: 20, loss: 0.15692676603794098
step: 30, loss: 0.10515639185905457
step: 40, loss: 0.02144516259431839
step: 50, loss: 0.10764260590076447
step: 60, loss: 0.04216141253709793
step: 70, loss: 0.021687153726816177
```

```

step: 80, loss: 0.013223519548773766
step: 90, loss: 0.01676744408905506
step: 100, loss: 0.006197718437761068
step: 110, loss: 0.10050424933433533
step: 120, loss: 0.003146782983094454
step: 130, loss: 0.0017813274171203375
step: 140, loss: 0.0016747668851166964
step: 150, loss: 0.02229686826467514
step: 160, loss: 0.07272855937480927
step: 170, loss: 0.0014064250281080604
step: 180, loss: 0.016249293461441994
step: 190, loss: 0.0018115765415132046
/usr/local/lib/python3.7/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=0.7804878048780487, f1=0.8319559228650137,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0787472352385521
step: 10, loss: 0.0026257894933223724
step: 20, loss: 0.0020388017874211073
step: 30, loss: 0.1274370700120926
step: 40, loss: 0.0006861822912469506
step: 50, loss: 0.00102017552126199
step: 60, loss: 0.0012811095220968127
step: 70, loss: 0.010284355841577053
step: 80, loss: 0.0002576157567091286
step: 90, loss: 0.010031024925410748
step: 100, loss: 0.03209471330046654
step: 110, loss: 0.0005473730270750821
step: 120, loss: 0.00034864956978708506
step: 130, loss: 0.008524986915290356
step: 140, loss: 0.007316168397665024
step: 150, loss: 0.0058969114907085896
step: 160, loss: 0.010681867599487305
step: 170, loss: 0.11750589311122894
step: 180, loss: 0.010059828869998455
step: 190, loss: 0.040114372968673706
/usr/local/lib/python3.7/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=0.7696202531645571, f1=0.7783505154639175,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.009820666164159775
step: 10, loss: 0.13492487370967865
step: 20, loss: 0.005688570439815521
step: 30, loss: 0.0006328211165964603
step: 40, loss: 0.00035711549571715295
step: 50, loss: 0.001871917163953185
step: 60, loss: 0.00160794158000499
step: 70, loss: 0.0002917067613452673
step: 80, loss: 0.002018687780946493
step: 90, loss: 0.0032091462053358555
step: 100, loss: 0.004095065873116255
step: 110, loss: 0.0028945605736225843
step: 120, loss: 0.03197278082370758
step: 130, loss: 0.0017782892100512981
step: 140, loss: 0.0007806222420185804
step: 150, loss: 0.0033182206097990274
step: 160, loss: 0.0008195765549317002
step: 170, loss: 0.00284725078381598
step: 180, loss: 0.002130824141204357
step: 190, loss: 0.012821988202631474
/usr/local/lib/python3.7/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=0.7720207253886011, f1=0.7803617571059432,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011551848147064447
step: 10, loss: 0.0006589835975319147
step: 20, loss: 0.001139806816354394
step: 30, loss: 0.0006953968550078571
step: 40, loss: 0.018318425863981247
step: 50, loss: 0.0059136259369552135
step: 60, loss: 0.00034705453435890377
step: 70, loss: 0.0004894059966318309
step: 80, loss: 0.0002386409614700824
step: 90, loss: 0.0016448753885924816
step: 100, loss: 0.0002094080700771883
step: 110, loss: 0.0020400627981871367
step: 120, loss: 0.07876306772232056
step: 130, loss: 0.0020348099060356617
step: 140, loss: 0.0732012391090393
step: 150, loss: 0.0945490151643753
step: 160, loss: 0.010322893969714642
step: 170, loss: 0.000823065871372819
step: 180, loss: 0.001282755401916802
step: 190, loss: 0.012327992357313633

```

```

/usr/local/lib/python3.7/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=0.7843137254901961, f1=0.8133704735376045,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000496394990477711
step: 10, loss: 0.003897612914443016
step: 20, loss: 0.0007758729625493288
step: 30, loss: 0.0040021538734436035
step: 40, loss: 0.00040541545604355633
step: 50, loss: 0.004107883665710688
step: 60, loss: 0.0077058179304003716
step: 70, loss: 0.000767529709264636
step: 80, loss: 0.0002465029829181731
step: 90, loss: 0.0001178368620458059
step: 100, loss: 0.00694124074652791
step: 110, loss: 0.00023470654559787363
step: 120, loss: 0.0002704511280171573
step: 130, loss: 0.0020691603422164917
step: 140, loss: 0.00033130444353446364
step: 150, loss: 0.00022390068625099957
step: 160, loss: 0.0019891206175088882
step: 170, loss: 0.00014251991524361074
step: 180, loss: 0.002291143173351884
step: 190, loss: 0.010583478026092052
/usr/local/lib/python3.7/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=0.79155672823219, f1=0.8074866310160429,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011895773059222847
step: 10, loss: 0.008260332979261875
step: 20, loss: 0.00020943333220202476
step: 30, loss: 0.00019653482013382018
step: 40, loss: 0.0013473685830831528
step: 50, loss: 0.0005141290603205562
step: 60, loss: 0.003715305356308818
step: 70, loss: 0.00012155337753938511
step: 80, loss: 0.00014821565127931535
step: 90, loss: 0.00011065641592722386
step: 100, loss: 0.00016515652532689273
step: 110, loss: 0.0002605137415230274

```

```

step: 120, loss: 0.00011150308273499832
step: 130, loss: 0.0019621264655143023
step: 140, loss: 0.05533221364021301
step: 150, loss: 0.00302359601482749
step: 160, loss: 6.909709918545559e-05
step: 170, loss: 0.0016791446832939982
step: 180, loss: 8.491877815686166e-05
step: 190, loss: 0.00027892139041796327
/usr/local/lib/python3.7/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=0.7819148936170213, f1=0.8074866310160429,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021661966457031667
step: 10, loss: 0.006949624978005886
step: 20, loss: 0.0006368350004777312
step: 30, loss: 0.0005057216621935368
step: 40, loss: 0.001504602376371622
step: 50, loss: 0.0004919660859741271
step: 60, loss: 0.00021544632909353822
step: 70, loss: 0.011352067813277245
step: 80, loss: 0.0014820432988926768
step: 90, loss: 0.0008096472593024373
step: 100, loss: 0.0003163034562021494
step: 110, loss: 0.001254456932656467
step: 120, loss: 0.002014310797676444
step: 130, loss: 0.00016114956815727055
step: 140, loss: 0.00015231952420435846
step: 150, loss: 0.0001448726252419874
step: 160, loss: 0.0002599644649308175
step: 170, loss: 0.00013737894187215716
step: 180, loss: 0.00010899094195337966
step: 190, loss: 0.0035384241491556168
/usr/local/lib/python3.7/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=0.7927461139896373, f1=0.8041775456919059,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0017575712408870459
step: 10, loss: 0.00010057335020974278
step: 20, loss: 0.00023732347472105175
step: 30, loss: 0.00013606117863673717

```

```

step: 40, loss: 0.005826808512210846
step: 50, loss: 0.00024187754024751484
step: 60, loss: 0.0024171287659555674
step: 70, loss: 0.00014441073290072381
step: 80, loss: 9.79569012997672e-05
step: 90, loss: 0.015694869682192802
step: 100, loss: 0.0011931498302146792
step: 110, loss: 0.0008432154427282512
step: 120, loss: 0.0006743901176378131
step: 130, loss: 0.00017030244634952396
step: 140, loss: 0.00013509979180525988
step: 150, loss: 0.000307488429825753
step: 160, loss: 0.003292315872386098
step: 170, loss: 0.0001792072580428794
step: 180, loss: 0.00037515279836952686
step: 190, loss: 0.00040341776912100613
/usr/local/lib/python3.7/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=0.7886597938144331, f1=0.8031914893617021,
best_f1=0.8067226890756303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030568061396479607
step: 10, loss: 0.00012478334247134626
step: 20, loss: 0.00027884787414222956
step: 30, loss: 7.549197471234947e-05
step: 40, loss: 0.00028444372583180666
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 50, loss: 0.0004198344540782273
step: 60, loss: 0.0005897119990549982
step: 70, loss: 0.00333242304623127
step: 80, loss: 0.00019406255159992725
step: 90, loss: 0.00015144811186473817
step: 100, loss: 0.0034373304806649685
step: 110, loss: 0.00027443678118288517
step: 120, loss: 0.00019577248895075172
step: 130, loss: 0.00011307690147077665
step: 140, loss: 0.0004192823253106326
step: 150, loss: 0.0004650400369428098
step: 160, loss: 0.00039087337790988386
step: 170, loss: 0.0001086026313714683
step: 180, loss: 0.0033409115858376026
step: 190, loss: 0.001903884345665574
/usr/local/lib/python3.7/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=0.7877237851662403, f1=0.7989556135770235,
best_f1=0.8067226890756303
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

```
2049it [00:08, 236.24it/s]
```

```
load_f1 = 0.7888888888888889
```

```
real_f1 = 0.7922437673130193
```

```
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, 269.45it/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 was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.6234651207923889
step: 10, loss: 0.3660033345222473
step: 20, loss: 0.3052654564380646
step: 30, loss: 0.3816978335380554
step: 40, loss: 0.2901850640773773
step: 50, loss: 0.23998615145683289
step: 60, loss: 0.24358031153678894
step: 70, loss: 0.359989732503891
step: 80, loss: 0.33422017097473145
step: 90, loss: 0.21354743838310242
step: 100, loss: 0.1895367056131363
step: 110, loss: 0.2970481514930725
step: 120, loss: 0.23095208406448364
step: 130, loss: 0.06381166726350784
step: 140, loss: 0.17095507681369781
step: 150, loss: 0.40349850058555603
step: 160, loss: 0.1441141664981842
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.27909743785858154
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6741071428571428, f1=0.6582278481012658,

```

```

best_f1=0.6582278481012658
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.16510754823684692
step: 10, loss: 0.14475712180137634
step: 20, loss: 0.12593019008636475
step: 30, loss: 0.3485892117023468
step: 40, loss: 0.11496519297361374
step: 50, loss: 0.022241033613681793
step: 60, loss: 0.065445676445961
step: 70, loss: 0.09229028224945068
step: 80, loss: 0.08351104706525803
step: 90, loss: 0.07766825705766678
step: 100, loss: 0.20404177904129028
step: 110, loss: 0.02346600592136383
step: 120, loss: 0.07156392931938171
step: 130, loss: 0.07365410774946213
step: 140, loss: 0.2606634497642517
step: 150, loss: 0.4033121168613434
step: 160, loss: 0.15904513001441956
step: 170, loss: 0.0524023175239563
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7536945812807883, f1=0.7547169811320755,
best_f1=0.7547169811320755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06392024457454681
step: 10, loss: 0.027947118505835533
step: 20, loss: 0.028427081182599068
step: 30, loss: 0.0853901132941246
step: 40, loss: 0.078388512134552
step: 50, loss: 0.08903764933347702
step: 60, loss: 0.07479194551706314
step: 70, loss: 0.15792442858219147
step: 80, loss: 0.07438082993030548
step: 90, loss: 0.16768231987953186
step: 100, loss: 0.004501659423112869
step: 110, loss: 0.06917589902877808
step: 120, loss: 0.09151157736778259
step: 130, loss: 0.1701078563928604

```

```

step: 140, loss: 0.04875453934073448
step: 150, loss: 0.04042021930217743
step: 160, loss: 0.044332873076200485
step: 170, loss: 0.21440911293029785
/usr/local/lib/python3.7/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=0.7397959183673468, f1=0.7462686567164178,
best_f1=0.7547169811320755
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01219121553003788
step: 10, loss: 0.18438521027565002
step: 20, loss: 0.005622598342597485
step: 30, loss: 0.07013305276632309
step: 40, loss: 0.0019618633668869734
step: 50, loss: 0.09663866460323334
step: 60, loss: 0.44163286685943604
step: 70, loss: 0.0038427787367254496
step: 80, loss: 0.1069841980934143
step: 90, loss: 0.06874772161245346
step: 100, loss: 0.18523643910884857
step: 110, loss: 0.06850843131542206
step: 120, loss: 0.004783527925610542
step: 130, loss: 0.04360024631023407
step: 140, loss: 0.07346707582473755
step: 150, loss: 0.04748911038041115
step: 160, loss: 0.1131085455417633
step: 170, loss: 0.04606867954134941
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7581047381546135, f1=0.800982800982801,
best_f1=0.800982800982801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006808717735111713
step: 10, loss: 0.02426321618258953
step: 20, loss: 0.01837329939007759
step: 30, loss: 0.13374827802181244
step: 40, loss: 0.02090870402753353
step: 50, loss: 0.14109748601913452

```

```

step: 60, loss: 0.07171876728534698
step: 70, loss: 0.3178672790527344
step: 80, loss: 0.07979930937290192
step: 90, loss: 0.07838863134384155
step: 100, loss: 0.0434732548892498
step: 110, loss: 0.09291849285364151
step: 120, loss: 0.03749977424740791
step: 130, loss: 0.04461924731731415
step: 140, loss: 0.06415316462516785
step: 150, loss: 0.02606600522994995
step: 160, loss: 0.054839469492435455
step: 170, loss: 0.020643463358283043
/usr/local/lib/python3.7/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=0.7399463806970509, f1=0.7792207792207791,
best_f1=0.800982800982801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016711367294192314
step: 10, loss: 0.0032185146119445562
step: 20, loss: 0.014074756763875484
step: 30, loss: 0.0033370519522577524
step: 40, loss: 0.00851406529545784
step: 50, loss: 0.10108137875795364
step: 60, loss: 0.044846758246421814
step: 70, loss: 0.13703197240829468
step: 80, loss: 0.19399313628673553
step: 90, loss: 0.029525576159358025
step: 100, loss: 0.015259717591106892
step: 110, loss: 0.0012821122072637081
step: 120, loss: 0.01214972510933876
step: 130, loss: 0.010121370665729046
step: 140, loss: 0.011471251025795937
step: 150, loss: 0.04914963245391846
step: 160, loss: 0.08880621194839478
step: 170, loss: 0.007943261414766312
/usr/local/lib/python3.7/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=0.7318181818181818, f1=0.7418655097613883,
best_f1=0.800982800982801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014860739465802908
step: 10, loss: 0.005053987726569176

```

```

step: 20, loss: 0.02337154559791088
step: 30, loss: 0.01364572811871767
step: 40, loss: 0.011544949375092983
step: 50, loss: 0.011757755652070045
step: 60, loss: 0.002221804577857256
step: 70, loss: 0.00029469977016560733
step: 80, loss: 0.02614353969693184
step: 90, loss: 0.0004565906128846109
step: 100, loss: 0.24261701107025146
step: 110, loss: 0.006531793624162674
step: 120, loss: 0.0018170720431953669
step: 130, loss: 0.13500452041625977
step: 140, loss: 0.002355113159865141
step: 150, loss: 0.006774050183594227
step: 160, loss: 0.00045016148942522705
step: 170, loss: 0.024367090314626694
/usr/local/lib/python3.7/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=0.7286063569682152, f1=0.7440758293838862,
best_f1=0.800982800982801
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03718109428882599
step: 10, loss: 0.0008889214368537068
step: 20, loss: 0.008922005072236061
step: 30, loss: 0.0399584136903286
step: 40, loss: 0.00027961324667558074
step: 50, loss: 0.0011420889059081674
step: 60, loss: 0.00612499937415123
step: 70, loss: 0.0005624795448966324
step: 80, loss: 0.007005623076111078
step: 90, loss: 0.0008221374009735882
step: 100, loss: 0.07037787139415741
step: 110, loss: 0.05091012641787529
step: 120, loss: 0.2472604364156723
step: 130, loss: 0.0028903144411742687
step: 140, loss: 0.0019419884774833918
step: 150, loss: 0.02345333620905876
step: 160, loss: 0.00764778396114707
step: 170, loss: 0.005398216191679239
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7594339622641509, f1=0.7586206896551724,
best_f1=0.7586206896551724
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006221872754395008
step: 10, loss: 0.0138468649238348
step: 20, loss: 0.0025925994850695133
step: 30, loss: 0.017405645921826363
step: 40, loss: 0.0030084867030382156
step: 50, loss: 0.00029376090969890356
step: 60, loss: 0.00566469132900238
step: 70, loss: 0.07225128263235092
step: 80, loss: 0.005776825360953808
step: 90, loss: 0.0362560898065567
step: 100, loss: 0.0010181987890973687
step: 110, loss: 0.0005575891118496656
step: 120, loss: 0.008736607618629932
step: 130, loss: 0.07679009437561035
step: 140, loss: 0.0007005081279203296
step: 150, loss: 0.0012759589590132236
step: 160, loss: 0.03539247810840607
step: 170, loss: 0.08338268846273422
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7634408602150538, f1=0.781725888324873,
best_f1=0.781725888324873
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05263035371899605
step: 10, loss: 0.004972298163920641
step: 20, loss: 0.027131337672472
step: 30, loss: 0.0008762114448472857
step: 40, loss: 0.011385058052837849
step: 50, loss: 0.03956514596939087
step: 60, loss: 0.002766585675999522
step: 70, loss: 0.002285279333591461
step: 80, loss: 0.006619642488658428
step: 90, loss: 0.011255850084125996
step: 100, loss: 0.0003723430563695729
step: 110, loss: 0.005823239218443632

```



```

step: 120, loss: 0.000856635335367173
step: 130, loss: 0.01514729019254446
step: 140, loss: 0.008625940419733524
step: 150, loss: 0.0055745383724570274
step: 160, loss: 0.004528796300292015
step: 170, loss: 0.0007149585289880633
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7666666666666666, f1=0.771764705882353,
best_f1=0.771764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03601359575986862
step: 10, loss: 0.02459665946662426
step: 20, loss: 0.012121854349970818
step: 30, loss: 0.00034163129748776555
step: 40, loss: 0.001837565447203815
step: 50, loss: 0.0035295295529067516
step: 60, loss: 0.01546016987413168
step: 70, loss: 0.00010267527250107378
step: 80, loss: 0.00012374226935207844
step: 90, loss: 0.00017265968199353665
step: 100, loss: 0.0001422372879460454
step: 110, loss: 0.016332248225808144
step: 120, loss: 0.002662380924448371
step: 130, loss: 0.0003295237256679684
step: 140, loss: 0.11095841228961945
step: 150, loss: 0.11507043242454529
step: 160, loss: 0.019585512578487396
step: 170, loss: 0.004997307434678078
/usr/local/lib/python3.7/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=0.7572815533980582, f1=0.7540983606557378,
best_f1=0.771764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014192926464602351
step: 10, loss: 0.0007118404610082507
step: 20, loss: 0.00046362928696908057
step: 30, loss: 0.0001835105795180425

```

```

step: 40, loss: 0.0002063263236777857
step: 50, loss: 0.00010488492989679798
step: 60, loss: 0.00017516470688860863
step: 70, loss: 0.11211751401424408
step: 80, loss: 0.00013388645311351866
step: 90, loss: 0.00048263900680467486
step: 100, loss: 0.0036939906422048807
step: 110, loss: 0.00041188058094121516
step: 120, loss: 0.17406325042247772
step: 130, loss: 0.0006461957818828523
step: 140, loss: 0.002347957342863083
step: 150, loss: 0.004322394262999296
step: 160, loss: 0.000948545231949538
step: 170, loss: 0.00027815226349048316
/usr/local/lib/python3.7/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=0.7506297229219144, f1=0.7745098039215688,
best_f1=0.771764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003950126003473997
step: 10, loss: 0.0011578170815482736
step: 20, loss: 0.11389105021953583
step: 30, loss: 0.00014665478374809027
step: 40, loss: 0.005711263045668602
step: 50, loss: 0.00019099605560768396
step: 60, loss: 0.0017486430006101727
step: 70, loss: 0.0008412525057792664
step: 80, loss: 0.0007300340221263468
step: 90, loss: 0.0001155482605099678
step: 100, loss: 0.00022515324235428125
step: 110, loss: 0.00010943094093818218
step: 120, loss: 0.007887208834290504
step: 130, loss: 0.00047874372103251517
step: 140, loss: 0.0003838582488242537
step: 150, loss: 0.00391613133251667
step: 160, loss: 0.001005985657684505
step: 170, loss: 0.002935280092060566
/usr/local/lib/python3.7/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=0.7575757575757577, f1=0.7853658536585365,
best_f1=0.771764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.00028780687716789544
step: 10, loss: 0.00011592009832384065
step: 20, loss: 0.00020596683316398412
step: 30, loss: 0.0002020819520112127
step: 40, loss: 0.00023810061975382268
step: 50, loss: 0.00023026936105452478
step: 60, loss: 0.00018723808170761913
step: 70, loss: 0.00023086011060513556
step: 80, loss: 0.000656842952594161
step: 90, loss: 0.0001281618169741705
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0029067588038742542
step: 110, loss: 0.0004798480949830264
step: 120, loss: 0.0014679552987217903
step: 130, loss: 0.0007649438921362162
step: 140, loss: 0.0003075620043091476
step: 150, loss: 8.468585292575881e-05
step: 160, loss: 0.00012308389705140144
step: 170, loss: 0.00042754338937811553
/usr/local/lib/python3.7/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=0.742857142857143, f1=0.7791563275434243,
best_f1=0.771764705882353
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004946084809489548
step: 10, loss: 0.0011764372466132045
step: 20, loss: 0.002080217469483614
step: 30, loss: 0.00031602371018379927
step: 40, loss: 0.002493385225534439
step: 50, loss: 0.00019321056606713682
step: 60, loss: 0.0019694326911121607
step: 70, loss: 9.231232252204791e-05
step: 80, loss: 0.0007332575041800737
step: 90, loss: 0.00019148230785503983
step: 100, loss: 0.041867852210998535
step: 110, loss: 0.00013775669503957033
step: 120, loss: 0.0017039697850123048
step: 130, loss: 0.010080907493829727
step: 140, loss: 0.0014891327591612935
step: 150, loss: 0.0003527305962052196
step: 160, loss: 0.0002127970365108922
step: 170, loss: 0.00022153090685606003
/usr/local/lib/python3.7/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=0.7450980392156863, f1=0.776595744680851,
best_f1=0.771764705882353
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, 267.33it/s]
```

```
load_f1 = 0.7673860911270983
```

```
real_f1 = 0.7677725118483412
```

```
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, 232.46it/s]
```

2.3 BASELINE DIRTY

2.3.1 DBLP-ACM - Running the matcher

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

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

```

cast_model_type      : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights       : True
loss_scale            : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.6441285014152527
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.6109216213226318
step: 20, loss: 0.4959302842617035
step: 30, loss: 0.305388867855072
step: 40, loss: 0.29025641083717346
step: 50, loss: 0.12691541016101837
step: 60, loss: 0.25182417035102844
step: 70, loss: 0.12026908993721008
step: 80, loss: 0.08253105729818344
step: 90, loss: 0.14339222013950348
step: 100, loss: 0.010541040450334549
step: 110, loss: 0.21868067979812622
step: 120, loss: 0.03393853083252907
step: 130, loss: 0.10634499788284302
step: 140, loss: 0.02894618548452854
step: 150, loss: 0.09402076154947281
step: 160, loss: 0.04380695894360542
step: 170, loss: 0.019305579364299774
step: 180, loss: 0.04629803076386452
step: 190, loss: 0.05862690880894661
step: 200, loss: 0.15840040147304535
step: 210, loss: 0.06474887579679489
step: 220, loss: 0.0303948987275362
step: 230, loss: 0.12013732641935349
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9597315436241611, f1=0.9561304836895389,
best_f1=0.9561304836895389
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006581639405339956
step: 10, loss: 0.02444346621632576
step: 20, loss: 0.3877164125442505
step: 30, loss: 0.04399847984313965
step: 40, loss: 0.07888227701187134
step: 50, loss: 0.01521238777864933
step: 60, loss: 0.014061784371733665
step: 70, loss: 0.06905391067266464
step: 80, loss: 0.007679387461394072
step: 90, loss: 0.0432644747197628
step: 100, loss: 0.018896058201789856
step: 110, loss: 0.07210248708724976
step: 120, loss: 0.09167592227458954
step: 130, loss: 0.055799294263124466
step: 140, loss: 0.019588235765695572
step: 150, loss: 0.035338543355464935
step: 160, loss: 0.12429646402597427
step: 170, loss: 0.0056664771400392056
step: 180, loss: 0.08816742151975632
step: 190, loss: 0.006213505752384663
step: 200, loss: 0.027011172845959663
step: 210, loss: 0.0071349297650158405
step: 220, loss: 0.05038592219352722
step: 230, loss: 0.010538686998188496
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9613259668508287, f1=0.9565217391304347,
best_f1=0.9565217391304347
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05183551087975502
step: 10, loss: 0.0073532527312636375

```

```

step: 20, loss: 0.0779389813542366
step: 30, loss: 0.01109229400753975
step: 40, loss: 0.12400711327791214
step: 50, loss: 0.053965386003255844
step: 60, loss: 0.018795805051922798
step: 70, loss: 0.027045149356126785
step: 80, loss: 0.013146914541721344
step: 90, loss: 0.00927499495446682
step: 100, loss: 0.005408504977822304
step: 110, loss: 0.001675393315963447
step: 120, loss: 0.026576122269034386
step: 130, loss: 0.0015247536357492208
step: 140, loss: 0.013556737452745438
step: 150, loss: 0.028170380741357803
step: 160, loss: 0.016569755971431732
step: 170, loss: 0.005048944614827633
step: 180, loss: 0.00506582111120224
step: 190, loss: 0.029078125953674316
step: 200, loss: 0.002852616598829627
step: 210, loss: 0.009671502746641636
step: 220, loss: 0.03163892775774002
step: 230, loss: 0.0019614044576883316
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9696287964004499, f1=0.9683972911963882,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002992290072143078
step: 10, loss: 0.011452151462435722
step: 20, loss: 0.0017823732923716307
step: 30, loss: 0.0007445430965162814
step: 40, loss: 0.023545093834400177
step: 50, loss: 0.000964981212746352
step: 60, loss: 0.020623652264475822
step: 70, loss: 0.028081053867936134
step: 80, loss: 0.004126721061766148
step: 90, loss: 0.008289230987429619
step: 100, loss: 0.0030258281622081995
step: 110, loss: 0.001063512871041894
step: 120, loss: 0.014775888994336128
step: 130, loss: 0.008534924127161503

```



```

step: 140, loss: 0.004071706440299749
step: 150, loss: 0.11078827828168869
step: 160, loss: 0.003531928174197674
step: 170, loss: 0.008117295801639557
step: 180, loss: 0.0005587866180576384
step: 190, loss: 0.0036432179622352123
step: 200, loss: 0.0023055539932101965
step: 210, loss: 0.004601523280143738
step: 220, loss: 0.0010733745293691754
step: 230, loss: 0.01672612689435482
/usr/local/lib/python3.7/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=0.9664429530201343, f1=0.967741935483871,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002645728411152959
step: 10, loss: 0.0981890857219696
step: 20, loss: 0.0019644177518785
step: 30, loss: 0.00853273831307888
step: 40, loss: 0.0006614226149395108
step: 50, loss: 0.0010820950847119093
step: 60, loss: 0.1385490447282791
step: 70, loss: 0.02902495674788952
step: 80, loss: 0.000605712819378823
step: 90, loss: 0.0010530742583796382
step: 100, loss: 0.0007243221043609083
step: 110, loss: 0.0005063890712335706
step: 120, loss: 0.00015138419985305518
step: 130, loss: 0.01114693284034729
step: 140, loss: 0.0006444664322771132
step: 150, loss: 0.023649778217077255
step: 160, loss: 0.0003393237420823425
step: 170, loss: 0.0008363024680875242
step: 180, loss: 0.0007109529688023031
step: 190, loss: 0.2134261280298233
step: 200, loss: 0.0047856541350483894
step: 210, loss: 0.000984686310403049
step: 220, loss: 0.0023001490626484156
step: 230, loss: 0.0008477923111058772
/usr/local/lib/python3.7/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=0.9657458563535912, f1=0.954895489548955,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003907154779881239
step: 10, loss: 0.0007890769629739225
step: 20, loss: 0.0018033962696790695
step: 30, loss: 0.0006103129708208144
step: 40, loss: 0.0002595445257611573
step: 50, loss: 0.0011723871575668454
step: 60, loss: 0.0002663711493369192
step: 70, loss: 0.0019463502103462815
step: 80, loss: 0.002356046810746193
step: 90, loss: 0.0013341910671442747
step: 100, loss: 0.004067517351359129
step: 110, loss: 0.002404573140665889
step: 120, loss: 0.0019287285394966602
step: 130, loss: 0.00025630276650190353
step: 140, loss: 0.00018497841665521264
step: 150, loss: 0.0005761755164712667
step: 160, loss: 0.0007608196465298533
step: 170, loss: 0.00022149710275698453
step: 180, loss: 0.011092865839600563
step: 190, loss: 0.008051086217164993
step: 200, loss: 0.00024962707539089024
step: 210, loss: 0.0005959462723694742
step: 220, loss: 0.0009968926897272468
step: 230, loss: 0.03883456066250801
/usr/local/lib/python3.7/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=0.9688195991091313, f1=0.9643652561247216,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002532112877815962
step: 10, loss: 0.0003529653185978532
step: 20, loss: 0.000301858177408576
step: 30, loss: 0.0004533208557404578
step: 40, loss: 0.002374847186729312
step: 50, loss: 0.0005343423690646887
step: 60, loss: 0.0004313638783060014
step: 70, loss: 0.0004896354512311518
step: 80, loss: 0.0008418255602009594
step: 90, loss: 0.0023588119074702263
step: 100, loss: 0.00040800878196023405
step: 110, loss: 0.026280928403139114
step: 120, loss: 0.00040508172241970897
step: 130, loss: 0.0001363862829748541

```

```

step: 140, loss: 0.0011369847925379872
step: 150, loss: 0.0009126010118052363
step: 160, loss: 0.033131301403045654
step: 170, loss: 0.01001206785440445
step: 180, loss: 0.0017610522918403149
step: 190, loss: 0.0004554844636004418
step: 200, loss: 0.013862443156540394
step: 210, loss: 0.00022463641653303057
step: 220, loss: 0.0004381001053843647
step: 230, loss: 0.00035338912857696414
/usr/local/lib/python3.7/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=0.9596412556053813, f1=0.9562289562289561,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021637449390254915
step: 10, loss: 0.0004878729523625225
step: 20, loss: 0.05729437619447708
step: 30, loss: 0.0022415448911488056
step: 40, loss: 0.009536578319966793
step: 50, loss: 0.0005912721389904618
step: 60, loss: 0.0001943773531820625
step: 70, loss: 0.00043002620805054903
step: 80, loss: 0.0004319207218941301
step: 90, loss: 0.0001775069977156818
step: 100, loss: 0.00016062681970652193
step: 110, loss: 0.000540036300662905
step: 120, loss: 0.004338833969086409
step: 130, loss: 0.0005393823958002031
step: 140, loss: 0.0029930558521300554
step: 150, loss: 0.00011765466479118913
step: 160, loss: 0.0005368213169276714
step: 170, loss: 0.00038687564665451646
step: 180, loss: 0.02079421654343605
step: 190, loss: 0.0007685561431571841
step: 200, loss: 0.0003082443727180362
step: 210, loss: 0.0007295868126675487
step: 220, loss: 0.00029179290868341923
step: 230, loss: 0.0002504356671124697
/usr/local/lib/python3.7/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=0.9642857142857144, f1=0.9608938547486034,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018132510012947023
step: 10, loss: 0.0003644790267571807
step: 20, loss: 0.00040490340325050056
step: 30, loss: 0.00013686400779988617
step: 40, loss: 0.04987344518303871
step: 50, loss: 0.0005689330282621086
step: 60, loss: 0.0002453587076161057
step: 70, loss: 0.0009130765683948994
step: 80, loss: 0.00020062307885382324
step: 90, loss: 0.000441663694800809
step: 100, loss: 0.0004454999871086329
step: 110, loss: 9.665799734648317e-05
step: 120, loss: 9.135348227573559e-05
step: 130, loss: 0.0001335780107183382
step: 140, loss: 9.737162326928228e-05
step: 150, loss: 0.013332456350326538
step: 160, loss: 0.00013515830505639315
step: 170, loss: 0.00015856184472795576
step: 180, loss: 0.00247013452462852
step: 190, loss: 0.00017068098532035947
step: 200, loss: 9.750050230650231e-05
step: 210, loss: 5.827046697959304e-05
step: 220, loss: 0.0001055106331477873
step: 230, loss: 0.0001484119420638308
/usr/local/lib/python3.7/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=0.9679558011049725, f1=0.9621380846325166,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.864873234415427e-05
step: 10, loss: 5.82476714043878e-05
step: 20, loss: 7.092941086739302e-05
step: 30, loss: 8.958144462667406e-05
step: 40, loss: 0.00011541189451236278
step: 50, loss: 0.0005322793149389327
step: 60, loss: 0.0001023961667669937
step: 70, loss: 0.001350433100014925
step: 80, loss: 0.000123656791402027
step: 90, loss: 0.00011586394248297438
step: 100, loss: 8.240182069130242e-05
step: 110, loss: 0.0008995484095066786
step: 120, loss: 0.00034192309249192476
step: 130, loss: 0.00012951833195984364

```

```

step: 140, loss: 0.03884255513548851
step: 150, loss: 0.03426220268011093
step: 160, loss: 8.996805263450369e-05
step: 170, loss: 0.00013409004895947874
step: 180, loss: 0.000586646725423634
step: 190, loss: 0.003156122984364629
step: 200, loss: 0.00011232511315029114
step: 210, loss: 0.00023139787663239986
step: 220, loss: 8.363651431864128e-05
step: 230, loss: 0.0003643095842562616
/usr/local/lib/python3.7/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=0.9656699889258028, f1=0.9588431590656283,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006669017020612955
step: 10, loss: 0.00017147143080364913
step: 20, loss: 0.0010471061104908586
step: 30, loss: 0.0008873838814906776
step: 40, loss: 0.00010161793761653826
step: 50, loss: 0.0001619568356545642
step: 60, loss: 0.0004946395056322217
step: 70, loss: 6.816424138378352e-05
step: 80, loss: 0.00020971171034034342
step: 90, loss: 9.495942504145205e-05
step: 100, loss: 0.00014697939332108945
step: 110, loss: 0.00012106277426937595
step: 120, loss: 5.472785414895043e-05
step: 130, loss: 3.466987982392311e-05
step: 140, loss: 5.359394344850443e-05
step: 150, loss: 0.03722933679819107
step: 160, loss: 4.9621768994256854e-05
step: 170, loss: 0.005903564393520355
step: 180, loss: 0.00013775346451438963
step: 190, loss: 5.452154437080026e-05
step: 200, loss: 9.922509343596175e-05
step: 210, loss: 3.954954809159972e-05
step: 220, loss: 6.796640082029626e-05
step: 230, loss: 7.199739047791809e-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=0.9615384615384616, f1=0.9522727272727274,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 5.797533958684653e-05
step: 10, loss: 3.800002014031634e-05
step: 20, loss: 0.00012199994671391323
step: 30, loss: 9.559416503179818e-05
step: 40, loss: 0.00010029796976596117
step: 50, loss: 0.0011065161088481545
step: 60, loss: 0.0003205709799658507
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 4.26821134169586e-05
step: 80, loss: 0.0004494424210861325
step: 90, loss: 4.492053994908929e-05
step: 100, loss: 3.9348120481008664e-05
step: 110, loss: 0.00010157944780075923
step: 120, loss: 5.484707435243763e-05
step: 130, loss: 5.5753782362444326e-05
step: 140, loss: 9.449598292121664e-05
step: 150, loss: 4.9428152124164626e-05
step: 160, loss: 5.621396121568978e-05
step: 170, loss: 4.585248461808078e-05
step: 180, loss: 0.00014772414579056203
step: 190, loss: 0.00018485462351236492
step: 200, loss: 3.0606035579694435e-05
step: 210, loss: 5.162590605323203e-05
step: 220, loss: 3.945959542761557e-05
step: 230, loss: 6.824937008786947e-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=0.9672316384180792, f1=0.9592760180995475,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00023807960678823292
step: 10, loss: 7.557478966191411e-05
step: 20, loss: 4.583743793773465e-05
step: 30, loss: 5.570462963078171e-05
step: 40, loss: 0.00015925892512314022
step: 50, loss: 0.00012097090802853927
step: 60, loss: 0.007685823831707239
step: 70, loss: 3.172685319441371e-05
step: 80, loss: 4.704781895270571e-05
step: 90, loss: 0.00025740868295542896
step: 100, loss: 2.9071139579173177e-05
step: 110, loss: 0.00047603031271137297
step: 120, loss: 0.015414129942655563

```

```

step: 130, loss: 0.00013136942288838327
step: 140, loss: 4.56363704870455e-05
step: 150, loss: 4.636657831724733e-05
step: 160, loss: 0.00017696956638246775
step: 170, loss: 0.0001619770482648164
step: 180, loss: 4.4362353946780786e-05
step: 190, loss: 3.692311656777747e-05
step: 200, loss: 4.459583942661993e-05
step: 210, loss: 2.6221616280963644e-05
step: 220, loss: 4.14454989368096e-05
step: 230, loss: 4.948806235915981e-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=0.9688888888888889, f1=0.9617977528089887,
best_f1=0.9683972911963882
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.938488953281194e-05
step: 10, loss: 0.000261232431512326
step: 20, loss: 4.5131484512239695e-05
step: 30, loss: 4.8772344598546624e-05
step: 40, loss: 0.0006730247405357659
step: 50, loss: 4.033954974147491e-05
step: 60, loss: 7.14636844350025e-05
step: 70, loss: 7.219616236397997e-05
step: 80, loss: 5.676085493178107e-05
step: 90, loss: 4.343774344306439e-05
step: 100, loss: 3.690856829052791e-05
step: 110, loss: 3.427137926337309e-05
step: 120, loss: 2.24258255911991e-05
step: 130, loss: 4.0008042560657486e-05
step: 140, loss: 4.377400910016149e-05
step: 150, loss: 2.9749058739980683e-05
step: 160, loss: 0.0012832671636715531
step: 170, loss: 1.9881577827618457e-05
step: 180, loss: 4.677494871430099e-05
step: 190, loss: 4.927003101329319e-05
step: 200, loss: 3.150729389744811e-05
step: 210, loss: 4.305124093662016e-05
step: 220, loss: 2.309608316863887e-05
step: 230, loss: 0.0006046998896636069
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9707207207207207, f1=0.9625425652667423,
best_f1=0.9625425652667423
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.77339913736796e-05
step: 10, loss: 2.461591429891996e-05
step: 20, loss: 6.0058169765397906e-05
step: 30, loss: 4.18702547904104e-05
step: 40, loss: 3.902839671354741e-05
step: 50, loss: 0.015735724940896034
step: 60, loss: 7.652047497685999e-05
step: 70, loss: 4.2849889723584056e-05
step: 80, loss: 4.198448732495308e-05
step: 90, loss: 0.0001337937719654292
step: 100, loss: 6.210917490534484e-05
step: 110, loss: 3.5314140404807404e-05
step: 120, loss: 5.4481792176375166e-05
step: 130, loss: 3.699060471262783e-05
step: 140, loss: 3.259150616941042e-05
step: 150, loss: 6.220671639312059e-05
step: 160, loss: 2.8467604352044873e-05
step: 170, loss: 2.733155997702852e-05
step: 180, loss: 6.65540574118495e-05
step: 190, loss: 5.794252501800656e-05
step: 200, loss: 5.066532321507111e-05
step: 210, loss: 3.578385440050624e-05
step: 220, loss: 9.2658432549797e-05
step: 230, loss: 3.893098619300872e-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=0.9696287964004499, f1=0.9626274065685164,
best_f1=0.9625425652667423
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```



```

Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:13, 178.43it/s]
load_f1 = 0.9706546275395034
real_f1 = 0.9685393258426966
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, 231.06it/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 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: 0.6434397101402283

step: 10, loss: 0.5378242135047913
step: 20, loss: 0.5508438944816589
step: 30, loss: 0.28909194469451904
step: 40, loss: 0.21086615324020386
step: 50, loss: 0.2620730996131897
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 60, loss: 0.14341111481189728
step: 70, loss: 0.1557777225971222
step: 80, loss: 0.18285831809043884
step: 90, loss: 0.1879182904958725
step: 100, loss: 0.055151667445898056
step: 110, loss: 0.15539368987083435
step: 120, loss: 0.1376592367887497
step: 130, loss: 0.08112350106239319
step: 140, loss: 0.09381970763206482
step: 150, loss: 0.051277074962854385
step: 160, loss: 0.02501608431339264
step: 170, loss: 0.330187052488327
step: 180, loss: 0.11588793247938156
step: 190, loss: 0.01837746426463127
step: 200, loss: 0.2165653258562088
step: 210, loss: 0.0893593356013298
step: 220, loss: 0.19313760101795197
step: 230, loss: 0.16423268616199493
step: 240, loss: 0.0740782618522644
step: 250, loss: 0.06980437785387039
step: 260, loss: 0.21048061549663544
step: 270, loss: 0.029024651274085045
step: 280, loss: 0.06965041160583496
step: 290, loss: 0.23357471823692322
step: 300, loss: 0.04537446051836014
step: 310, loss: 0.2443067580461502
step: 320, loss: 0.1446043699979782
step: 330, loss: 0.05266736447811127
step: 340, loss: 0.03287513926625252
step: 350, loss: 0.1286092847585678
step: 360, loss: 0.07144635170698166
step: 370, loss: 0.1582929491996765
step: 380, loss: 0.030854441225528717
step: 390, loss: 0.2832420766353607
step: 400, loss: 0.29775795340538025
step: 410, loss: 0.06783635169267654
step: 420, loss: 0.10455572605133057
step: 430, loss: 0.11124283820390701
step: 440, loss: 0.024779126048088074
step: 450, loss: 0.02290944568812847
step: 460, loss: 0.011413751170039177
step: 470, loss: 0.1442476063966751

```

step: 480, loss: 0.05772489309310913
step: 490, loss: 0.07790610194206238
step: 500, loss: 0.048180751502513885
step: 510, loss: 0.059100158512592316
step: 520, loss: 0.08379337936639786
step: 530, loss: 0.01876796782016754
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9232183908045977, f1=0.9137931034482759,
best_f1=0.9137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11643245816230774
step: 10, loss: 0.18706384301185608
step: 20, loss: 0.06453488767147064
step: 30, loss: 0.023183535784482956
step: 40, loss: 0.08931440114974976
step: 50, loss: 0.17925409972667694
step: 60, loss: 0.01874520443379879
step: 70, loss: 0.03319424390792847
step: 80, loss: 0.16972891986370087
step: 90, loss: 0.013384317979216576
step: 100, loss: 0.07410163432359695
step: 110, loss: 0.0215121041983366
step: 120, loss: 0.13092255592346191
step: 130, loss: 0.05324448272585869
step: 140, loss: 0.0648813247680664
step: 150, loss: 0.07819919288158417
step: 160, loss: 0.10278784483671188
step: 170, loss: 0.08617507666349411
step: 180, loss: 0.0518556609749794
step: 190, loss: 0.029959242790937424
step: 200, loss: 0.03967658057808876
step: 210, loss: 0.06988416612148285
step: 220, loss: 0.057042255997657776
step: 230, loss: 0.011354928836226463
step: 240, loss: 0.05498243123292923
step: 250, loss: 0.01615300588309765
step: 260, loss: 0.00349560659378767
step: 270, loss: 0.2735513150691986
step: 280, loss: 0.16380757093429565
step: 290, loss: 0.02332318015396595

```

```

step: 300, loss: 0.15154637396335602
step: 310, loss: 0.008965893648564816
step: 320, loss: 0.20111201703548431
step: 330, loss: 0.07879937440156937
step: 340, loss: 0.027778789401054382
step: 350, loss: 0.005369293969124556
step: 360, loss: 0.04881337657570839
step: 370, loss: 0.15888598561286926
step: 380, loss: 0.043742358684539795
step: 390, loss: 0.11084030568599701
step: 400, loss: 0.04354004189372063
step: 410, loss: 0.021875489503145218
step: 420, loss: 0.14340856671333313
step: 430, loss: 0.02579553984105587
step: 440, loss: 0.10004041343927383
step: 450, loss: 0.052376072853803635
step: 460, loss: 0.06698812544345856
step: 470, loss: 0.1000409722328186
step: 480, loss: 0.17168405652046204
step: 490, loss: 0.014373158104717731
step: 500, loss: 0.14914470911026
step: 510, loss: 0.03501846268773079
step: 520, loss: 0.04284599423408508
step: 530, loss: 0.05514281988143921
/usr/local/lib/python3.7/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=0.919294990723562, f1=0.9196470041802136,
best_f1=0.9137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08807898312807083
step: 10, loss: 0.04845432937145233
step: 20, loss: 0.06142435595393181
step: 30, loss: 0.08439323306083679
step: 40, loss: 0.11271840333938599
step: 50, loss: 0.1403663009405136
step: 60, loss: 0.034616172313690186
step: 70, loss: 0.015391386114060879
step: 80, loss: 0.0007827401859685779
step: 90, loss: 0.004634593613445759
step: 100, loss: 0.01270271185785532
step: 110, loss: 0.0037941315677016973
step: 120, loss: 0.00476818298920989
step: 130, loss: 0.0030742797534912825
step: 140, loss: 0.01700134389102459
step: 150, loss: 0.06923218816518784

```

```

step: 160, loss: 0.01930665783584118
step: 170, loss: 0.15370449423789978
step: 180, loss: 0.019712327048182487
step: 190, loss: 0.0167235117405653
step: 200, loss: 0.005654735956341028
step: 210, loss: 0.04914185777306557
step: 220, loss: 0.2285894900560379
step: 230, loss: 0.07005481421947479
step: 240, loss: 0.004330806899815798
step: 250, loss: 0.026661256328225136
step: 260, loss: 0.020218422636389732
step: 270, loss: 0.00918429996818304
step: 280, loss: 0.2775265574455261
step: 290, loss: 0.007971461862325668
step: 300, loss: 0.018958164379000664
step: 310, loss: 0.003982999362051487
step: 320, loss: 0.00605979235842824
step: 330, loss: 0.015651021152734756
step: 340, loss: 0.025043407455086708
step: 350, loss: 0.158258855342865
step: 360, loss: 0.04484665393829346
step: 370, loss: 0.030256621539592743
step: 380, loss: 0.0497741661965847
step: 390, loss: 0.024465026333928108
step: 400, loss: 0.02100278250873089
step: 410, loss: 0.029015399515628815
step: 420, loss: 0.12255379557609558
step: 430, loss: 0.05273621529340744
step: 440, loss: 0.007167628034949303
step: 450, loss: 0.08829446882009506
step: 460, loss: 0.10493530333042145
step: 470, loss: 0.02916329726576805
step: 480, loss: 0.008696830831468105
step: 490, loss: 0.006268684286624193
step: 500, loss: 0.2611404061317444
step: 510, loss: 0.03518218919634819
step: 520, loss: 0.05674657225608826
step: 530, loss: 0.09271793067455292
/usr/local/lib/python3.7/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=0.9217311233885819, f1=0.9164764947512551,
best_f1=0.9137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004849723540246487
step: 10, loss: 0.004482378251850605

```

step: 20, loss: 0.0023784360382705927
step: 30, loss: 0.003966504707932472
step: 40, loss: 0.011802173219621181
step: 50, loss: 0.03138786181807518
step: 60, loss: 0.00416486244648695
step: 70, loss: 0.007561716716736555
step: 80, loss: 0.0021248222328722477
step: 90, loss: 0.1090255007147789
step: 100, loss: 0.0040401918813586235
step: 110, loss: 0.10962199419736862
step: 120, loss: 0.10746313631534576
step: 130, loss: 0.0022417374420911074
step: 140, loss: 0.0018838284304365516
step: 150, loss: 0.02501172386109829
step: 160, loss: 0.03605464845895767
step: 170, loss: 0.005214938428252935
step: 180, loss: 0.001498937956057489
step: 190, loss: 0.031077628955245018
step: 200, loss: 0.06001976132392883
step: 210, loss: 0.029787492007017136
step: 220, loss: 0.010995608754456043
step: 230, loss: 0.1798762083053589
step: 240, loss: 0.018070697784423828
step: 250, loss: 0.019923623651266098
step: 260, loss: 0.006976708769798279
step: 270, loss: 0.008429375477135181
step: 280, loss: 0.025221319869160652
step: 290, loss: 0.016496527940034866
step: 300, loss: 0.0003466905909590423
step: 310, loss: 0.0025122826918959618
step: 320, loss: 0.001906127668917179
step: 330, loss: 0.020006088539958
step: 340, loss: 0.004930358380079269
step: 350, loss: 0.007200541906058788
step: 360, loss: 0.11077626794576645
step: 370, loss: 0.009054267778992653
step: 380, loss: 0.0052681565284729
step: 390, loss: 0.009950417093932629
step: 400, loss: 0.0022019040770828724
step: 410, loss: 0.0044242385774850845
step: 420, loss: 0.007756626233458519
step: 430, loss: 0.0483374185860157
step: 440, loss: 0.08385147154331207
step: 450, loss: 0.0030037774704396725
step: 460, loss: 0.023379158228635788
step: 470, loss: 0.009564675390720367
step: 480, loss: 0.0392158105969429
step: 490, loss: 0.005089115351438522

```

step: 500, loss: 0.0023274302948266268
step: 510, loss: 0.003044974058866501
step: 520, loss: 0.01811772584915161
step: 530, loss: 0.04035509005188942
/usr/local/lib/python3.7/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=0.9128668171557561, f1=0.9168173598553345,
best_f1=0.9137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01880708523094654
step: 10, loss: 0.02581530250608921
step: 20, loss: 0.0035921777598559856
step: 30, loss: 0.13204224407672882
step: 40, loss: 0.0864294171333313
step: 50, loss: 0.041495371609926224
step: 60, loss: 0.095211923122406
step: 70, loss: 0.01595485210418701
step: 80, loss: 0.004986106883734465
step: 90, loss: 0.0026693374384194613
step: 100, loss: 0.0025271927006542683
step: 110, loss: 0.012039121240377426
step: 120, loss: 0.049634162336587906
step: 130, loss: 0.01793290115892887
step: 140, loss: 0.0007306007901206613
step: 150, loss: 0.001811023335903883
step: 160, loss: 0.0018053185194730759
step: 170, loss: 0.023408744484186172
step: 180, loss: 0.0018712312448769808
step: 190, loss: 0.0013584501575678587
step: 200, loss: 0.0007407473167404532
step: 210, loss: 0.006662196945399046
step: 220, loss: 0.0014263689517974854
step: 230, loss: 0.00785746704787016
step: 240, loss: 0.0012481800513342023
step: 250, loss: 0.3117237687110901
step: 260, loss: 0.13881923258304596
step: 270, loss: 0.04206959530711174
step: 280, loss: 0.006306846160441637
step: 290, loss: 0.17208965122699738
step: 300, loss: 0.007663773372769356
step: 310, loss: 0.0003859448479488492
step: 320, loss: 0.16664516925811768
step: 330, loss: 0.007465972099453211
step: 340, loss: 0.0022199642844498158
step: 350, loss: 0.005567531101405621

```



```

step: 360, loss: 0.006516800262033939
step: 370, loss: 0.01271621510386467
step: 380, loss: 0.0004207420570310205
step: 390, loss: 0.0019113158341497183
step: 400, loss: 0.01554343942552805
step: 410, loss: 0.0008141085272654891
step: 420, loss: 0.0019825417548418045
step: 430, loss: 0.01252529863268137
step: 440, loss: 0.004201109521090984
step: 450, loss: 0.0014982790453359485
step: 460, loss: 0.004111847374588251
step: 470, loss: 0.004733523819595575
step: 480, loss: 0.0020604762248694897
step: 490, loss: 0.002246921183541417
step: 500, loss: 0.00657288171350956
step: 510, loss: 0.0043153720907866955
step: 520, loss: 0.012928865849971771
step: 530, loss: 0.002197843976318836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9232927970065481, f1=0.9250814332247558,
best_f1=0.9250814332247558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004198269918560982
step: 10, loss: 0.00016825196507852525
step: 20, loss: 0.14010384678840637
step: 30, loss: 0.0009831845527514815
step: 40, loss: 0.0022861582692712545
step: 50, loss: 0.016566231846809387
step: 60, loss: 0.0019863820634782314
step: 70, loss: 0.005214978475123644
step: 80, loss: 0.001682356814853847
step: 90, loss: 0.004578023683279753
step: 100, loss: 0.012183954939246178
step: 110, loss: 0.0007094950997270644
step: 120, loss: 0.0003271047316957265
step: 130, loss: 0.061979904770851135
step: 140, loss: 0.031225129961967468
step: 150, loss: 0.050861574709415436
step: 160, loss: 0.0009602899081073701
step: 170, loss: 0.004831794183701277

```

```

step: 180, loss: 0.015880059450864792
step: 190, loss: 0.0011158377164974809
step: 200, loss: 0.006729897577315569
step: 210, loss: 0.019763333722949028
step: 220, loss: 0.0006661724764853716
step: 230, loss: 0.001259073382243514
step: 240, loss: 0.07754389196634293
step: 250, loss: 0.0009777318919077516
step: 260, loss: 0.00027012432110495865
step: 270, loss: 0.0722898542881012
step: 280, loss: 0.0012736795470118523
step: 290, loss: 0.0004438654868863523
step: 300, loss: 0.0012775621144101024
step: 310, loss: 0.00439996924251318
step: 320, loss: 0.002834339626133442
step: 330, loss: 0.0004000192566309124
step: 340, loss: 0.05304238200187683
step: 350, loss: 0.0014265027130022645
step: 360, loss: 0.0028706397861242294
step: 370, loss: 0.0004222370916977525
step: 380, loss: 0.00024659818154759705
step: 390, loss: 0.01398205105215311
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 400, loss: 0.12133575230836868
step: 410, loss: 0.0002739351475611329
step: 420, loss: 0.0023963383864611387
step: 430, loss: 0.000160722220627591
step: 440, loss: 0.00011240437743254006
step: 450, loss: 0.0005305737140588462
step: 460, loss: 0.0001624017022550106
step: 470, loss: 0.0014962460845708847
step: 480, loss: 0.02202105149626732
step: 490, loss: 0.000531012483406812
step: 500, loss: 0.00016741118452046067
step: 510, loss: 0.0019818011205643415
step: 520, loss: 0.0011465736897662282
step: 530, loss: 0.04273118078708649
/usr/local/lib/python3.7/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=0.919020715630885, f1=0.9163160355306218,
best_f1=0.9250814332247558
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006490717642009258
step: 10, loss: 0.00019227559096179903
step: 20, loss: 0.00043063017074018717

```

step: 30, loss: 0.0004940208746120334
step: 40, loss: 0.10041859745979309
step: 50, loss: 0.0008059705141931772
step: 60, loss: 0.009416261687874794
step: 70, loss: 0.03966149315237999
step: 80, loss: 0.005597769282758236
step: 90, loss: 0.00028744604787789285
step: 100, loss: 0.002864591544494033
step: 110, loss: 0.0004603659035637975
step: 120, loss: 0.00026390518178232014
step: 130, loss: 0.09393730759620667
step: 140, loss: 0.00018242117948830128
step: 150, loss: 0.000206462835194543
step: 160, loss: 8.959163824329153e-05
step: 170, loss: 0.0006346361478790641
step: 180, loss: 0.0005089825135655701
step: 190, loss: 0.0001663690054556355
step: 200, loss: 0.00010269471385981888
step: 210, loss: 0.11960452049970627
step: 220, loss: 8.126216562232003e-05
step: 230, loss: 0.01789427362382412
step: 240, loss: 0.000853582052513957
step: 250, loss: 0.0005548386834561825
step: 260, loss: 0.0011360736098140478
step: 270, loss: 0.028539977967739105
step: 280, loss: 7.512732554459944e-05
step: 290, loss: 0.00017657178977970034
step: 300, loss: 0.1305568665266037
step: 310, loss: 0.0004130998859182
step: 320, loss: 7.308850763365626e-05
step: 330, loss: 0.0014526807935908437
step: 340, loss: 0.0811847522854805
step: 350, loss: 0.004070157650858164
step: 360, loss: 0.009612586349248886
step: 370, loss: 0.0007204993162304163
step: 380, loss: 0.005558141507208347
step: 390, loss: 0.0014100114349275827
step: 400, loss: 0.0102480323985219
step: 410, loss: 0.004183376673609018
step: 420, loss: 0.0006334320059977472
step: 430, loss: 0.0003245319821871817
step: 440, loss: 0.0002876554208341986
step: 450, loss: 0.0002943317813333124
step: 460, loss: 0.00047092378372326493
step: 470, loss: 0.0022038938477635384
step: 480, loss: 0.07758425921201706
step: 490, loss: 0.02232929691672325
step: 500, loss: 0.07263181358575821

```

step: 510, loss: 0.004903942812234163
step: 520, loss: 0.01858852431178093
step: 530, loss: 0.0013654881622642279
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9236396890717878, f1=0.9231468849477036,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003718295774888247
step: 10, loss: 0.029670067131519318
step: 20, loss: 0.00010634477075655013
step: 30, loss: 0.03313319385051727
step: 40, loss: 0.005824750289320946
step: 50, loss: 0.006026541348546743
step: 60, loss: 0.005793341901153326
step: 70, loss: 0.003976090345531702
step: 80, loss: 0.0008985574822872877
step: 90, loss: 0.00034448530641384423
step: 100, loss: 0.0005802613450214267
step: 110, loss: 0.0002602569875307381
step: 120, loss: 0.0003268627915531397
step: 130, loss: 0.008222161792218685
step: 140, loss: 0.0009181100176647305
step: 150, loss: 6.828139157732949e-05
step: 160, loss: 0.0005342381773516536
step: 170, loss: 0.000677023665048182
step: 180, loss: 0.00011273905693087727
step: 190, loss: 0.0007620507385581732
step: 200, loss: 0.00033274738234467804
step: 210, loss: 0.0024355670902878046
step: 220, loss: 0.0011876297648996115
step: 230, loss: 0.03070945106446743
step: 240, loss: 0.00013891799608245492
step: 250, loss: 0.0001098791544791311
step: 260, loss: 0.0002696355804800987
step: 270, loss: 0.0050420318730175495
step: 280, loss: 0.006277227774262428
step: 290, loss: 0.014461299404501915
step: 300, loss: 0.0016836802242323756
step: 310, loss: 0.028160016983747482
step: 320, loss: 0.021506115794181824

```

```

step: 330, loss: 0.009891824796795845
step: 340, loss: 0.00033065120805986226
step: 350, loss: 0.04839123785495758
step: 360, loss: 0.09852293878793716
step: 370, loss: 9.909514483297244e-05
step: 380, loss: 0.006015400402247906
step: 390, loss: 0.0006547014345414937
step: 400, loss: 0.00013194164785090834
step: 410, loss: 4.271962461643852e-05
step: 420, loss: 0.000317638274282217
step: 430, loss: 0.06139913946390152
step: 440, loss: 0.0059999278746545315
step: 450, loss: 0.0004621703992597759
step: 460, loss: 0.043028607964515686
step: 470, loss: 8.750735287321731e-05
step: 480, loss: 0.000520825560670346
step: 490, loss: 0.09439483284950256
step: 500, loss: 0.010577445849776268
step: 510, loss: 0.008959801867604256
step: 520, loss: 0.0014344804221764207
step: 530, loss: 0.0006896275444887578
/usr/local/lib/python3.7/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=0.9181692094313454, f1=0.9224376731301939,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001232940354384482
step: 10, loss: 7.346767233684659e-05
step: 20, loss: 0.00042749333078972995
step: 30, loss: 0.00011704774806275964
step: 40, loss: 4.4310341763775796e-05
step: 50, loss: 0.0002583555760793388
step: 60, loss: 5.509636321221478e-05
step: 70, loss: 0.00026263235486112535
step: 80, loss: 0.003359345719218254
step: 90, loss: 0.00011629951040958986
step: 100, loss: 4.9768234021030366e-05
step: 110, loss: 4.230555714457296e-05
step: 120, loss: 0.00027219587354920805
step: 130, loss: 0.00047947486746124923
step: 140, loss: 0.000109684988274239
step: 150, loss: 0.00010741678124759346
step: 160, loss: 4.438921678229235e-05
step: 170, loss: 0.0011335494928061962
step: 180, loss: 4.316862759878859e-05

```

```

step: 190, loss: 0.00011336867464706302
step: 200, loss: 0.00012515668640844524
step: 210, loss: 0.00011878268560394645
step: 220, loss: 0.0007942196680232882
step: 230, loss: 0.00025534184533171356
step: 240, loss: 5.77219580009114e-05
step: 250, loss: 6.474983092630282e-05
step: 260, loss: 0.00134507124312222
step: 270, loss: 8.817190973786637e-05
step: 280, loss: 0.009471439756453037
step: 290, loss: 0.04795824736356735
step: 300, loss: 0.000116143943159841
step: 310, loss: 0.00240780645981431
step: 320, loss: 0.00020073342602699995
step: 330, loss: 0.0003068959340453148
step: 340, loss: 0.00013237394159659743
step: 350, loss: 0.005845887120813131
step: 360, loss: 0.00018065865151584148
step: 370, loss: 0.00027557279099710286
step: 380, loss: 4.592612094711512e-05
step: 390, loss: 4.348819129518233e-05
step: 400, loss: 0.003839269280433655
step: 410, loss: 0.0019656952936202288
step: 420, loss: 0.00014962356362957507
step: 430, loss: 0.0033976042177528143
step: 440, loss: 0.0026058319490402937
step: 450, loss: 0.0001159970197477378
step: 460, loss: 0.00021887969342060387
step: 470, loss: 0.0011943718418478966
step: 480, loss: 0.00012835334928240627
step: 490, loss: 0.0005985737661831081
step: 500, loss: 0.000883443804923445
step: 510, loss: 0.00037753849755972624
step: 520, loss: 0.0003666893462650478
step: 530, loss: 0.00017523397400509566
/usr/local/lib/python3.7/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=0.9103840682788051, f1=0.9168241965973536,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00025038898456841707
step: 10, loss: 4.004327274742536e-05
step: 20, loss: 2.7577219952945597e-05
step: 30, loss: 0.0001514641917310655
step: 40, loss: 0.00024406693410128355

```

step: 50, loss: 0.00037805380998179317
step: 60, loss: 0.00018643811927177012
step: 70, loss: 0.0014055358478799462
step: 80, loss: 4.0681730752112344e-05
step: 90, loss: 0.0006535246502608061
step: 100, loss: 4.4073884055251256e-05
step: 110, loss: 2.3293319827644154e-05
step: 120, loss: 3.8479916838696226e-05
step: 130, loss: 0.00039254711009562016
step: 140, loss: 0.0165173951536417
step: 150, loss: 0.0004340809246059507
step: 160, loss: 3.7973568396409974e-05
step: 170, loss: 0.0002370672591496259
step: 180, loss: 6.921061867615208e-05
step: 190, loss: 0.0003845598257612437
step: 200, loss: 0.0013988139107823372
step: 210, loss: 0.0007424675859510899
step: 220, loss: 0.0001311264350079
step: 230, loss: 0.011216796934604645
step: 240, loss: 0.006070106755942106
step: 250, loss: 4.392546907183714e-05
step: 260, loss: 0.012644169852137566
step: 270, loss: 0.00010385709902038798
step: 280, loss: 0.00023266306379809976
step: 290, loss: 5.734811566071585e-05
step: 300, loss: 0.00011531866039149463
step: 310, loss: 0.00011836691555799916
step: 320, loss: 0.0034520362969487906
step: 330, loss: 4.935732067679055e-05
step: 340, loss: 0.04121209681034088
step: 350, loss: 0.00012945827620569617
step: 360, loss: 0.0005559457931667566
step: 370, loss: 0.00031842160387896
step: 380, loss: 0.00039124811883084476
step: 390, loss: 0.013088593259453773
step: 400, loss: 0.003422756213694811
step: 410, loss: 0.0029302246402949095
step: 420, loss: 0.0009665403049439192
step: 430, loss: 0.0008826549747027457
step: 440, loss: 0.007330482825636864
step: 450, loss: 9.822897118283436e-05
step: 460, loss: 8.266422810265794e-05
step: 470, loss: 8.824786345940083e-05
step: 480, loss: 6.733728514518589e-05
step: 490, loss: 0.00013591547030955553
step: 500, loss: 0.1356464922428131
step: 510, loss: 0.00011034258204745129
step: 520, loss: 5.2176721510477364e-05

```
step: 530, loss: 0.00044527294812723994
/usr/local/lib/python3.7/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=0.9216589861751152, f1=0.9223659889094269,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003629823913797736
step: 10, loss: 0.00018700752116274089
step: 20, loss: 8.048838208196685e-05
step: 30, loss: 0.0010410496033728123
step: 40, loss: 0.0006278094951994717
step: 50, loss: 0.00020958969253115356
step: 60, loss: 0.02921128459274769
step: 70, loss: 0.0017346561653539538
step: 80, loss: 3.121322879451327e-05
step: 90, loss: 0.000707640778273344
step: 100, loss: 0.00043858433491550386
step: 110, loss: 0.003184923203662038
step: 120, loss: 4.956566044711508e-05
step: 130, loss: 9.019537537824363e-05
step: 140, loss: 2.854238118743524e-05
step: 150, loss: 2.6861478545470163e-05
step: 160, loss: 3.3063868613680825e-05
step: 170, loss: 3.6352550523588434e-05
step: 180, loss: 4.900147905573249e-05
step: 190, loss: 0.005470622330904007
step: 200, loss: 0.03180735558271408
step: 210, loss: 3.913544423994608e-05
step: 220, loss: 0.0022400261368602514
step: 230, loss: 0.00037624480319209397
step: 240, loss: 8.391439041588455e-05
step: 250, loss: 2.1684551029466093e-05
step: 260, loss: 3.579024996724911e-05
step: 270, loss: 0.0001905894750962034
step: 280, loss: 0.00031342709553427994
step: 290, loss: 0.005814043805003166
step: 300, loss: 0.0005973337101750076
step: 310, loss: 0.0060964603908360004
step: 320, loss: 9.393696382176131e-05
step: 330, loss: 0.00012555862485896796
step: 340, loss: 0.0002787297125905752
step: 350, loss: 3.810791531577706e-05
step: 360, loss: 3.30754483002238e-05
step: 370, loss: 0.00018864471348933876
step: 380, loss: 4.037213511765003e-05
```



```

step: 390, loss: 2.8221718821441755e-05
step: 400, loss: 5.26429976162035e-05
step: 410, loss: 2.2444552087108605e-05
step: 420, loss: 5.771850919700228e-05
step: 430, loss: 9.664794197306037e-05
step: 440, loss: 0.00020506879081949592
step: 450, loss: 2.809530997183174e-05
step: 460, loss: 0.017859792336821556
step: 470, loss: 4.951663868268952e-05
step: 480, loss: 4.7418594476766884e-05
step: 490, loss: 0.0001897506444947794
step: 500, loss: 0.0012201642384752631
step: 510, loss: 2.609420334920287e-05
step: 520, loss: 4.742216333397664e-05
step: 530, loss: 2.862768633349333e-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=0.9091755938518863, f1=0.9169776119402985,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018088123761117458
step: 10, loss: 2.9942189939902164e-05
step: 20, loss: 2.9253356842673384e-05
step: 30, loss: 0.002546164905652404
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 40, loss: 0.00019100680947303772
step: 50, loss: 0.0016163423424586654
step: 60, loss: 0.00013465674419421703
step: 70, loss: 2.8095482775825076e-05
step: 80, loss: 1.8398915926809423e-05
step: 90, loss: 0.0005142384325154126
step: 100, loss: 3.6572208045981824e-05
step: 110, loss: 2.0615476387320086e-05
step: 120, loss: 2.4712306185392663e-05
step: 130, loss: 2.65339531324571e-05
step: 140, loss: 0.0002573409001342952
step: 150, loss: 3.646667755674571e-05
step: 160, loss: 1.9728717234102078e-05
step: 170, loss: 3.377182292751968e-05
step: 180, loss: 4.019852713099681e-05
step: 190, loss: 0.00015212022117339075
step: 200, loss: 2.285795744683128e-05
step: 210, loss: 0.0008453777409158647
step: 220, loss: 5.414510451373644e-05
step: 230, loss: 0.001060683629475534

```

```

step: 240, loss: 0.00013354157272260636
step: 250, loss: 0.012730028480291367
step: 260, loss: 0.0001834847789723426
step: 270, loss: 9.8405871540308e-05
step: 280, loss: 2.2160971639095806e-05
step: 290, loss: 3.965984797105193e-05
step: 300, loss: 0.0003374536754563451
step: 310, loss: 7.322720193769783e-05
step: 320, loss: 0.00010163435945287347
step: 330, loss: 5.621254604193382e-05
step: 340, loss: 6.499156734207645e-05
step: 350, loss: 9.63422644417733e-05
step: 360, loss: 0.0001616410881979391
step: 370, loss: 0.00157643249258399
step: 380, loss: 0.0012952563120052218
step: 390, loss: 7.106162229320034e-05
step: 400, loss: 4.935550896334462e-05
step: 410, loss: 0.0004548759898170829
step: 420, loss: 0.0033378254156559706
step: 430, loss: 0.015128914266824722
step: 440, loss: 0.001229125540703535
step: 450, loss: 5.732976205763407e-05
step: 460, loss: 3.608084807638079e-05
step: 470, loss: 3.955667489208281e-05
step: 480, loss: 0.006396490149199963
step: 490, loss: 0.00030901809805072844
step: 500, loss: 0.0003682984388433397
step: 510, loss: 0.004941081162542105
step: 520, loss: 0.006390131078660488
step: 530, loss: 0.0004957308992743492
/usr/local/lib/python3.7/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=0.9151765245300322, f1=0.91804788213628,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0016309495549649
step: 10, loss: 8.997219265438616e-05
step: 20, loss: 0.0008246239158324897
step: 30, loss: 6.109229434514418e-05
step: 40, loss: 2.3386808607028797e-05
step: 50, loss: 0.0006854504463262856
step: 60, loss: 0.0005071907653473318
step: 70, loss: 5.1717976020881906e-05
step: 80, loss: 5.040510222897865e-05
step: 90, loss: 0.015562167391180992

```

```
step: 100, loss: 0.00011964979057665914
step: 110, loss: 0.00014837415073998272
step: 120, loss: 2.3282671463675797e-05
step: 130, loss: 2.065262924588751e-05
step: 140, loss: 0.00015084927144926041
step: 150, loss: 9.447153570363298e-05
step: 160, loss: 0.00037116603925824165
step: 170, loss: 6.266349373618141e-05
step: 180, loss: 6.333143392112106e-05
step: 190, loss: 0.0003725734131876379
step: 200, loss: 0.00013189055607654154
step: 210, loss: 4.4694195821648464e-05
step: 220, loss: 0.0004326516354922205
step: 230, loss: 2.153910645574797e-05
step: 240, loss: 0.006293728947639465
step: 250, loss: 3.6717923649121076e-05
step: 260, loss: 3.937500878237188e-05
step: 270, loss: 1.7478803783887997e-05
step: 280, loss: 2.5431872927583754e-05
step: 290, loss: 1.9963285012636334e-05
step: 300, loss: 4.2985375330317765e-05
step: 310, loss: 0.0027297933120280504
step: 320, loss: 0.005433185957372189
step: 330, loss: 0.0001565029815537855
step: 340, loss: 2.5755087335710414e-05
step: 350, loss: 2.2883945348439738e-05
step: 360, loss: 2.2373717001755722e-05
step: 370, loss: 0.0010255263186991215
step: 380, loss: 0.00012357017840258777
step: 390, loss: 0.00014457864745054394
step: 400, loss: 0.0003351573832333088
step: 410, loss: 0.0002740335476119071
step: 420, loss: 3.67562570318114e-05
step: 430, loss: 0.00011349433043505996
step: 440, loss: 6.632595614064485e-05
step: 450, loss: 5.205048364587128e-05
step: 460, loss: 0.0004598810337483883
step: 470, loss: 0.00010223518620477989
step: 480, loss: 4.490119681577198e-05
step: 490, loss: 3.605077654356137e-05
step: 500, loss: 2.719748044910375e-05
step: 510, loss: 3.300422758911736e-05
step: 520, loss: 0.0006749233580194414
step: 530, loss: 2.5480465410510078e-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=0.9130636913063691, f1=0.9164722351843211,
```

```
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.4538701584097e-05
step: 10, loss: 2.4526809283997864e-05
step: 20, loss: 9.465856419410557e-05
step: 30, loss: 3.3700514904921874e-05
step: 40, loss: 0.00010776826093206182
step: 50, loss: 0.0003126589290332049
step: 60, loss: 1.86485249287216e-05
step: 70, loss: 0.000616741250269115
step: 80, loss: 4.258881017449312e-05
step: 90, loss: 3.7191202864050865e-05
step: 100, loss: 0.00023063791741151363
step: 110, loss: 2.657191907928791e-05
step: 120, loss: 3.3328797144349664e-05
step: 130, loss: 0.0007593858754262328
step: 140, loss: 3.992296478827484e-05
step: 150, loss: 0.0002620806044433266
step: 160, loss: 2.8437858418328688e-05
step: 170, loss: 0.0005727573879994452
step: 180, loss: 0.00021595723228529096
step: 190, loss: 3.960693356930278e-05
step: 200, loss: 3.775977893383242e-05
step: 210, loss: 4.370357055449858e-05
step: 220, loss: 2.3938160666148178e-05
step: 230, loss: 0.00013756075350102037
step: 240, loss: 2.060801489278674e-05
step: 250, loss: 3.0147670258884318e-05
step: 260, loss: 5.4047432058723643e-05
step: 270, loss: 4.467583858058788e-05
step: 280, loss: 1.9166365746059455e-05
step: 290, loss: 4.49307553935796e-05
step: 300, loss: 5.554275048780255e-05
step: 310, loss: 2.27833788812859e-05
step: 320, loss: 6.0259553720243275e-05
step: 330, loss: 3.955247302656062e-05
step: 340, loss: 4.6534270950360224e-05
step: 350, loss: 0.00011351909051882103
step: 360, loss: 0.0012671987060457468
step: 370, loss: 2.7759366275859065e-05
step: 380, loss: 0.00016154476907104254
step: 390, loss: 0.030920447781682014
step: 400, loss: 3.430789729463868e-05
step: 410, loss: 0.00013246967864688486
step: 420, loss: 3.524188286974095e-05
step: 430, loss: 3.9869482861831784e-05
```

```
step: 440, loss: 5.827058703289367e-05
step: 450, loss: 0.000261332985246554
step: 460, loss: 2.4198963728849776e-05
step: 470, loss: 4.751637243316509e-05
step: 480, loss: 3.3421201806049794e-05
step: 490, loss: 2.29771358135622e-05
step: 500, loss: 2.8362497687339783e-05
step: 510, loss: 0.00010350728553021327
step: 520, loss: 3.6456349334912375e-05
step: 530, loss: 4.123570033698343e-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=0.9148148148148147, f1=0.9213691026827011,
best_f1=0.9231468849477036
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.005639622453600168
step: 10, loss: 0.00013023878273088485
step: 20, loss: 2.2444346541306004e-05
step: 30, loss: 0.00039045122684910893
step: 40, loss: 0.000691338733304292
step: 50, loss: 0.0007553871837444603
step: 60, loss: 1.8995013306266628e-05
step: 70, loss: 0.00043409812496975064
step: 80, loss: 4.0181304939324036e-05
step: 90, loss: 2.1781379473395646e-05
step: 100, loss: 0.000704100530128926
step: 110, loss: 2.558094638516195e-05
step: 120, loss: 2.761025280051399e-05
step: 130, loss: 0.01740448549389839
step: 140, loss: 2.017956467170734e-05
step: 150, loss: 2.6061092285090126e-05
step: 160, loss: 5.745997987105511e-05
step: 170, loss: 0.00033976894337683916
step: 180, loss: 1.7873726392281242e-05
step: 190, loss: 2.485829645593185e-05
step: 200, loss: 0.00011557290417840704
step: 210, loss: 4.923213418805972e-05
step: 220, loss: 2.7938544008065946e-05
step: 230, loss: 3.673899846035056e-05
step: 240, loss: 2.184459663112648e-05
step: 250, loss: 3.5682565794559196e-05
step: 260, loss: 1.8514443581807427e-05
step: 270, loss: 4.142355828662403e-05
step: 280, loss: 3.264495171606541e-05
step: 290, loss: 1.8015292880591005e-05
```

```

step: 300, loss: 9.739780944073573e-05
step: 310, loss: 3.171192656736821e-05
step: 320, loss: 2.2767819245927967e-05
step: 330, loss: 2.5688948880997486e-05
step: 340, loss: 1.9401000827201642e-05
step: 350, loss: 1.6525180399185047e-05
step: 360, loss: 0.0002249307290185243
step: 370, loss: 1.4811621440458111e-05
step: 380, loss: 2.4668313926667906e-05
step: 390, loss: 6.325276626739651e-05
step: 400, loss: 0.00016562930250074714
step: 410, loss: 3.7437814171426e-05
step: 420, loss: 0.0003343595308251679
step: 430, loss: 0.00010487224062671885
step: 440, loss: 3.5811797715723515e-05
step: 450, loss: 1.7050439055310562e-05
step: 460, loss: 1.728139613987878e-05
step: 470, loss: 3.7783152947667986e-05
step: 480, loss: 1.6480467820656486e-05
step: 490, loss: 2.1110776287969202e-05
step: 500, loss: 2.134538590325974e-05
step: 510, loss: 3.0441011404036544e-05
step: 520, loss: 2.2448215531767346e-05
step: 530, loss: 2.5941792046069168e-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=0.912739150723285, f1=0.9197761194029851,
best_f1=0.9231468849477036
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, 230.70it/s]
load_f1 = 0.9268738574040218
real_f1 = 0.9201661282879556
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, 240.61it/s]

```

3 DK

```

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

```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-
wheels/public/simple/
Collecting en_core_web_lg==2.2.5
  Downloading https://github.com/explosion/spacy-
models/releases/download/en_core_web_lg-2.2.5/en_core_web_lg-2.2.5.tar.gz (827.9
MB)
    |                                     | 827.9 MB 83.1 MB/s
Requirement already satisfied: spacy>=2.2.2 in
/usr/local/lib/python3.7/dist-packages (from en_core_web_lg==2.2.5) (2.2.3)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in
/usr/local/lib/python3.7/dist-packages (from
spacy>=2.2.2->en_core_web_lg==2.2.5) (2.0.6)
Requirement already satisfied: 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: 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: 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: thinc<7.4.0,>=7.3.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (7.3.1)

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

Requirement already satisfied: 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: setuptools in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (57.4.0)

Requirement already satisfied: srsly<1.1.0,>=0.1.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (1.0.5)

Requirement already satisfied: blis<0.5.0,>=0.4.0 in /usr/local/lib/python3.7/dist-packages (from spacy>=2.2.2->en_core_web_lg==2.2.5) (0.4.1)

Requirement already satisfied: 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: importlib-metadata>=0.20 in /usr/local/lib/python3.7/dist-packages (from catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.12.0)

Requirement already satisfied: typing-extensions>=3.6.4 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (4.1.1)

Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.7/dist-packages (from importlib-metadata>=0.20->catalogue<1.1.0,>=0.0.7->spacy>=2.2.2->en_core_web_lg==2.2.5) (3.8.1)

Requirement already satisfied: 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: 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: idna<3,>=2.5 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (2.10)

Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.7/dist-packages (from requests<3.0.0,>=2.13.0->spacy>=2.2.2->en_core_web_lg==2.2.5) (1.25.4)

Requirement already satisfied: 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=62477ee1a058321f1b20e81b3850c232336b7b944764cdf30195b3094370e073
  Stored in directory: /tmp/pip-ephem-wheel-cache-y_4au5fk/wheels/11/95/ba/2c36cc368c0bd339b44a791c2c1881a1fb714b78c29a4cb8f5
Successfully built en-core-web-lg
Installing collected packages: en-core-web-lg
Successfully installed en-core-web-lg-2.2.5
  Download and installation successful
You can now load the model via spacy.load('en_core_web_lg')

```

3.1 DK STRUCTURED

3.1.1 Beer - Running the matcher

```

[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
  --task Structured/Beer \
  --batch_size 32 \
  --max_len 256 \
  --lr 3e-5 \
  --n_epochs 15 \
  --finetuning \
  --lm 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 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: 0.5633316040039062

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

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

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

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=0.423076923076923, f1=0.4444444444444444,

```

best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5009447932243347
/usr/local/lib/python3.7/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=0.3880597014925373, f1=0.42857142857142855,
best_f1=0.4444444444444444
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.6335073113441467
/usr/local/lib/python3.7/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=0.39285714285714285, f1=0.43478260869565216,
best_f1=0.4444444444444444
/usr/local/lib/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 16384.0
step: 0, loss: 0.4546206295490265
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 4: dev_f1=0.6285714285714286, f1=0.39999999999999997,
best_f1=0.39999999999999997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.36902734637260437
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6875000000000001, f1=0.4137931034482759,
best_f1=0.4137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4033195376396179
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.56, f1=0.5, best_f1=0.4137931034482759
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.357076495885849
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7058823529411764, f1=0.5405405405405405,
best_f1=0.5405405405405405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06275928765535355
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7647058823529412, f1=0.5294117647058824,
best_f1=0.5294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08164987713098526
/usr/local/lib/python3.7/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=0.7647058823529412, f1=0.689655172413793,
best_f1=0.5294117647058824
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.047943536192178726
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 10: dev_f1=0.7857142857142857, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01596740260720253
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7999999999999999, f1=0.5925925925925927,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01036353874951601
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6666666666666666,
best_f1=0.5925925925925927
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02195422537624836
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.8275862068965518, f1=0.6666666666666666,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010301643051207066
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6923076923076924,
best_f1=0.6666666666666666
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008771195076406002
/usr/local/lib/python3.7/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=0.7999999999999999, f1=0.6923076923076924,
best_f1=0.6666666666666666
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
91it [00:00, 112923.57it/s]
/usr/local/lib/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 = 0.8125000000000001
real_f1 = 0.8235294117647058
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

4398it [00:22, 197.90it/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: 0.6055591702461243
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.5852863192558289
step: 20, loss: 0.454665869474411
step: 30, loss: 0.3573714792728424
step: 40, loss: 0.20722141861915588
step: 50, loss: 0.025822214782238007
step: 60, loss: 0.03861261159181595
step: 70, loss: 0.16726341843605042
step: 80, loss: 0.10879834741353989
step: 90, loss: 0.14676155149936676
step: 100, loss: 0.007978096604347229
step: 110, loss: 0.09027930349111557
step: 120, loss: 0.008005037903785706
step: 130, loss: 0.07899255305528641
step: 140, loss: 0.0035501988604664803
step: 150, loss: 0.18242545425891876
step: 160, loss: 0.04731030762195587
step: 170, loss: 0.16509166359901428
step: 180, loss: 0.13779063522815704
step: 190, loss: 0.06580109894275665
step: 200, loss: 0.030287783592939377

```



```

step: 210, loss: 0.007262904662638903
step: 220, loss: 0.020601186901330948
step: 230, loss: 0.03957098349928856
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9774774774774775, f1=0.9688195991091313,
best_f1=0.9688195991091313
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007579081691801548
step: 10, loss: 0.0030141756869852543
step: 20, loss: 0.1023208349943161
step: 30, loss: 0.17158812284469604
step: 40, loss: 0.04657228663563728
step: 50, loss: 0.01158084161579609
step: 60, loss: 0.002663396066054702
step: 70, loss: 0.1757393181324005
step: 80, loss: 0.01007208600640297
step: 90, loss: 0.14558203518390656
step: 100, loss: 0.05346151068806648
step: 110, loss: 0.0636150911450386
step: 120, loss: 0.10839402675628662
step: 130, loss: 0.053406935185194016
step: 140, loss: 0.0018055529799312353
step: 150, loss: 0.06165565922856331
step: 160, loss: 0.008991433307528496
step: 170, loss: 0.008564339950680733
step: 180, loss: 0.005043931771069765
step: 190, loss: 0.09418854862451553
step: 200, loss: 0.0016451262636110187
step: 210, loss: 0.0011603685561567545
step: 220, loss: 0.10147646069526672
step: 230, loss: 0.04993358999490738
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.977728285077951, 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: 0.011733615770936012
step: 10, loss: 0.0036573426332324743
step: 20, loss: 0.0010944780660793185
step: 30, loss: 0.0515265092253685
step: 40, loss: 0.10680070519447327
step: 50, loss: 0.003712903708219528
step: 60, loss: 0.11637851595878601
step: 70, loss: 0.0027670543640851974
step: 80, loss: 0.0013706773752346635
step: 90, loss: 0.028028061613440514
step: 100, loss: 0.004516322165727615
step: 110, loss: 0.0005310337874107063
step: 120, loss: 0.004335614386945963
step: 130, loss: 0.0039062106516212225
step: 140, loss: 0.016710884869098663
step: 150, loss: 0.0010372984688729048
step: 160, loss: 0.04700840637087822
step: 170, loss: 0.0034742930438369513
step: 180, loss: 0.015604253858327866
step: 190, loss: 0.001740228501148522
step: 200, loss: 0.03255200386047363
step: 210, loss: 0.000547505565918982
step: 220, loss: 0.0003407473850529641
step: 230, loss: 0.0005453316261991858
/usr/local/lib/python3.7/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=0.975, f1=0.9625425652667423, 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: 0.00023915874771773815
step: 10, loss: 0.0005011584726162255
step: 20, loss: 0.0005161755252629519
step: 30, loss: 0.0010341200977563858
step: 40, loss: 0.0076985531486570835
step: 50, loss: 0.001397031475789845
step: 60, loss: 0.0184420607984066
step: 70, loss: 0.0004724173923023045
step: 80, loss: 0.00041587563464418054
step: 90, loss: 0.017065279185771942
step: 100, loss: 0.0003661029040813446
step: 110, loss: 0.00034144226810894907
step: 120, loss: 0.0032880380749702454
step: 130, loss: 0.016055583953857422

```

```

step: 140, loss: 0.002790022175759077
step: 150, loss: 0.14646948873996735
step: 160, loss: 0.0013623839477077127
step: 170, loss: 0.006577782332897186
step: 180, loss: 0.0006034629186615348
step: 190, loss: 0.0047114151529967785
step: 200, loss: 0.00046504041529260576
step: 210, loss: 0.06941074132919312
step: 220, loss: 0.00044915222679264843
step: 230, loss: 0.06174978241324425
/usr/local/lib/python3.7/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=0.9771689497716894, f1=0.9761634506242906,
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: 0.004730393644422293
step: 10, loss: 0.0012942003086209297
step: 20, loss: 0.0012684406246989965
step: 30, loss: 0.0004096674092579633
step: 40, loss: 0.00038797827437520027
step: 50, loss: 0.00027478020638227463
step: 60, loss: 0.00036905601155012846
step: 70, loss: 0.00014388565614353865
step: 80, loss: 0.00016197661170735955
step: 90, loss: 0.0003225479740649462
step: 100, loss: 0.0003884286852553487
step: 110, loss: 0.0002618397120386362
step: 120, loss: 9.015649266075343e-05
step: 130, loss: 0.000720409385394305
step: 140, loss: 0.0015755366766825318
step: 150, loss: 0.00010506659600650892
step: 160, loss: 0.00034737412352114916
step: 170, loss: 0.000579283747356385
step: 180, loss: 0.0007977465866133571
step: 190, loss: 0.0010968631831929088
step: 200, loss: 0.005084012169390917
step: 210, loss: 0.0003079433226957917
step: 220, loss: 0.0020095654763281345
step: 230, loss: 0.0017903036205098033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9832402234636871, f1=0.9755555555555556,
best_f1=0.9755555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004256174899637699
step: 10, loss: 0.00041093354229815304
step: 20, loss: 0.043142642825841904
step: 30, loss: 0.003404435934498906
step: 40, loss: 0.00046823505545035005
step: 50, loss: 0.0007905569509603083
step: 60, loss: 0.00023590629280079156
step: 70, loss: 0.006043487228453159
step: 80, loss: 0.004146998282521963
step: 90, loss: 0.000492514984216541
step: 100, loss: 0.05531885847449303
step: 110, loss: 0.0020854424219578505
step: 120, loss: 0.010133427567780018
step: 130, loss: 0.0009283495601266623
step: 140, loss: 0.0002475869841873646
step: 150, loss: 0.0006435008253902197
step: 160, loss: 0.19604937732219696
step: 170, loss: 0.0056843943893909454
step: 180, loss: 0.004572867415845394
step: 190, loss: 0.014902906492352486
step: 200, loss: 0.0008877007639966905
step: 210, loss: 0.0007033881265670061
step: 220, loss: 0.00031349711935035884
step: 230, loss: 0.006228654179722071
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 6: dev_f1=0.9809203142536477, f1=0.9787234042553192,
best_f1=0.9755555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005255151190795004
step: 10, loss: 0.0002334881282877177
step: 20, loss: 0.0002645023923832923
step: 30, loss: 0.0003831123758573085
step: 40, loss: 0.00020710818353109062
step: 50, loss: 0.00010692428622860461
step: 60, loss: 0.00020009076979476959
step: 70, loss: 0.0005452879122458398
step: 80, loss: 0.0006875034305267036
step: 90, loss: 0.00020116163068450987

```

```

step: 100, loss: 0.00026281282771378756
step: 110, loss: 0.025745224207639694
step: 120, loss: 0.007178918924182653
step: 130, loss: 0.000338055775500834
step: 140, loss: 0.0006752467015758157
step: 150, loss: 0.00016820286691654474
step: 160, loss: 0.016375236213207245
step: 170, loss: 0.014085136353969574
step: 180, loss: 0.002314559882506728
step: 190, loss: 0.0002086028689518571
step: 200, loss: 0.04364656284451485
step: 210, loss: 8.333216101163998e-05
step: 220, loss: 0.0004702533478848636
step: 230, loss: 0.014432416297495365
/usr/local/lib/python3.7/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=0.9775280898876404, f1=0.9788182831661093,
best_f1=0.9755555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004350001981947571
step: 10, loss: 0.0004767899517901242
step: 20, loss: 0.00010109033610206097
step: 30, loss: 0.0003835110692307353
step: 40, loss: 0.00020936198416166008
step: 50, loss: 0.00011552366777323186
step: 60, loss: 0.010409562848508358
step: 70, loss: 0.0011659780284389853
step: 80, loss: 0.0017969643231481314
step: 90, loss: 0.00020523136481642723
step: 100, loss: 0.002508903155103326
step: 110, loss: 0.0004043311928398907
step: 120, loss: 0.006769102998077869
step: 130, loss: 0.0008587163174524903
step: 140, loss: 0.0001996508362935856
step: 150, loss: 0.00038868660340085626
step: 160, loss: 0.00036929413909092546
step: 170, loss: 9.068537474377081e-05
step: 180, loss: 0.00029200356220826507
step: 190, loss: 0.0002668771194294095
step: 200, loss: 0.00013441329065244645
step: 210, loss: 0.00018455601821187884
step: 220, loss: 0.0003412798105273396
step: 230, loss: 0.00018683556118048728
/usr/local/lib/python3.7/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=0.9796380090497738, f1=0.9842696629213483,
best_f1=0.9755555555555556
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.923019777284935e-05
step: 10, loss: 0.00034609783324413
step: 20, loss: 0.00021286441187839955
step: 30, loss: 5.733142461394891e-05
step: 40, loss: 0.00028014928102493286
step: 50, loss: 7.96669046394527e-05
step: 60, loss: 0.00017369927081745118
step: 70, loss: 0.0006317768711596727
step: 80, loss: 0.00013084657257422805
step: 90, loss: 0.00013084657257422805
step: 100, loss: 0.00041396712185814977
step: 110, loss: 5.038766539655626e-05
step: 120, loss: 6.820580165367573e-05
step: 130, loss: 0.0001970106241060421
step: 140, loss: 0.0002912391792051494
step: 150, loss: 0.0015094360569491982
step: 160, loss: 0.00011692170664900914
step: 170, loss: 0.0001311389059992507
step: 180, loss: 0.0077492413111031055
step: 190, loss: 0.005877101793885231
step: 200, loss: 7.906121027190238e-05
step: 210, loss: 0.002270861528813839
step: 220, loss: 0.00010900259803747758
step: 230, loss: 0.00018289685249328613
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841269841269841, f1=0.9853768278965129,
best_f1=0.9853768278965129
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011667631042655557
step: 10, loss: 4.579444066621363e-05
step: 20, loss: 9.46272411965765e-05
step: 30, loss: 0.0029729041270911694
step: 40, loss: 7.262279541464522e-05
step: 50, loss: 0.00023935602803248912
step: 60, loss: 0.000375115021597594

```

```

step: 70, loss: 0.00034313034848310053
step: 80, loss: 6.386646418832242e-05
step: 90, loss: 0.00034265968133695424
step: 100, loss: 0.00011812376760644838
step: 110, loss: 0.006791003979742527
step: 120, loss: 9.747875446919352e-05
step: 130, loss: 6.374280928866938e-05
step: 140, loss: 0.0004829768731724471
step: 150, loss: 0.0005071352934464812
step: 160, loss: 4.9777016101870686e-05
step: 170, loss: 4.727518535219133e-05
step: 180, loss: 0.00010075936006614938
step: 190, loss: 0.00017739862960297614
step: 200, loss: 5.29286298842635e-05
step: 210, loss: 4.012325734947808e-05
step: 220, loss: 0.0001633658102946356
step: 230, loss: 0.00027923492598347366
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9841986455981941, f1=0.9865168539325843,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.540679896716028e-05
step: 10, loss: 0.0003059861483052373
step: 20, loss: 9.368098108097911e-05
step: 30, loss: 0.0002037174126598984
step: 40, loss: 0.00012565402721520513
step: 50, loss: 9.387926547788084e-05
step: 60, loss: 8.189101936295629e-05
step: 70, loss: 9.774988575372845e-05
step: 80, loss: 0.000752945605199784
step: 90, loss: 5.9076577599626034e-05
step: 100, loss: 4.632482887245715e-05
step: 110, loss: 0.00024517549900338054
step: 120, loss: 4.070086288265884e-05
step: 130, loss: 4.485599129111506e-05
step: 140, loss: 6.434762326534837e-05
step: 150, loss: 6.509339436888695e-05
step: 160, loss: 4.661051571019925e-05
step: 170, loss: 7.23503835615702e-05
step: 180, loss: 5.5666281696176156e-05

```

```

step: 190, loss: 4.3761621782323346e-05
step: 200, loss: 9.43678169278428e-05
step: 210, loss: 4.8191286623477936e-05
step: 220, loss: 5.539829726330936e-05
step: 230, loss: 6.692943861708045e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9864253393665158, f1=0.9865168539325843,
best_f1=0.9865168539325843
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011602149606915191
step: 10, loss: 0.00012402806896716356
step: 20, loss: 5.177422281121835e-05
step: 30, loss: 0.0002925751614384353
step: 40, loss: 5.9694684750866145e-05
step: 50, loss: 0.00013581149687524885
step: 60, loss: 5.2787123422604054e-05
step: 70, loss: 5.036059155827388e-05
step: 80, loss: 4.112540773348883e-05
step: 90, loss: 3.408854900044389e-05
step: 100, loss: 3.33287644025404e-05
step: 110, loss: 4.9054720875574276e-05
step: 120, loss: 0.00012745067942887545
step: 130, loss: 3.493456097203307e-05
step: 140, loss: 0.009222407825291157
step: 150, loss: 5.101411443320103e-05
step: 160, loss: 4.6359156840480864e-05
step: 170, loss: 4.0224076656159014e-05
step: 180, loss: 0.00021235646272543818
step: 190, loss: 3.601492426241748e-05
step: 200, loss: 6.097215009503998e-05
step: 210, loss: 5.954681182629429e-05
step: 220, loss: 3.987228774349205e-05
step: 230, loss: 0.05112675204873085
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9864559819413092, f1=0.9854096520763187,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011620075383689255
step: 10, loss: 0.0001545722916489467
step: 20, loss: 6.720466626575217e-05
step: 30, loss: 0.0003755600191652775
step: 40, loss: 4.7412588173756376e-05
step: 50, loss: 5.9369202062953264e-05
step: 60, loss: 0.00014272346743382514
step: 70, loss: 2.8959608243894763e-05
step: 80, loss: 3.57466415152885e-05
step: 90, loss: 0.0001681749999988824
step: 100, loss: 2.4832284907461144e-05
step: 110, loss: 6.935890269232914e-05
step: 120, loss: 0.0026731661055237055
step: 130, loss: 4.504382377490401e-05
step: 140, loss: 3.7735069781774655e-05
step: 150, loss: 3.5419016057858244e-05
step: 160, loss: 3.69305198546499e-05
step: 170, loss: 0.05706063285470009
step: 180, loss: 3.8570135075133294e-05
step: 190, loss: 3.970802208641544e-05
step: 200, loss: 0.0001661783317103982
step: 210, loss: 4.754496330861002e-05
step: 220, loss: 4.1440849599894136e-05
step: 230, loss: 3.816004755208269e-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=0.984090909090909, f1=0.9864864864864865,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.97391922888346e-05
step: 10, loss: 6.414781091734767e-05
step: 20, loss: 5.3355030104285106e-05
step: 30, loss: 8.327376417582855e-05
step: 40, loss: 6.772499909857288e-05
step: 50, loss: 5.031728869653307e-05
step: 60, loss: 3.849890345009044e-05
step: 70, loss: 5.862631951458752e-05
step: 80, loss: 4.310319127398543e-05
step: 90, loss: 7.534661563113332e-05
step: 100, loss: 4.641398481908254e-05

```

```

step: 110, loss: 4.542419264907949e-05
step: 120, loss: 2.222090733994264e-05
step: 130, loss: 3.683054819703102e-05
step: 140, loss: 8.825497206998989e-05
step: 150, loss: 7.776134589221328e-05
step: 160, loss: 3.000253127538599e-05
step: 170, loss: 2.386728192504961e-05
step: 180, loss: 0.00010600915993563831
step: 190, loss: 4.47107340733055e-05
step: 200, loss: 3.9383379771607e-05
step: 210, loss: 4.277205152902752e-05
step: 220, loss: 4.240070120431483e-05
step: 230, loss: 3.469258081167936e-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=0.9852774631936579, f1=0.9853768278965129,
best_f1=0.9854096520763187
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.0171998989535496e-05
step: 10, loss: 2.6891930247074924e-05
step: 20, loss: 4.3973999709123746e-05
step: 30, loss: 4.8996316763805225e-05
step: 40, loss: 9.17761935852468e-05
step: 50, loss: 0.0002851650060620159
step: 60, loss: 4.9039961595553905e-05
step: 70, loss: 0.005219289567321539
step: 80, loss: 3.586556340451352e-05
step: 90, loss: 0.0006977440789341927
step: 100, loss: 3.431243749218993e-05
step: 110, loss: 0.03527999296784401
step: 120, loss: 9.186115494230762e-05
step: 130, loss: 8.10051933513023e-05
step: 140, loss: 2.9037846616120078e-05
step: 150, loss: 0.00040804955642670393
step: 160, loss: 2.666487307578791e-05
step: 170, loss: 2.4321834644069895e-05
step: 180, loss: 4.708948836196214e-05
step: 190, loss: 4.489426646614447e-05
step: 200, loss: 3.5389231925364584e-05
step: 210, loss: 3.3086820621974766e-05
step: 220, loss: 4.709099084720947e-05
step: 230, loss: 2.9682436434086412e-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=0.9852774631936579, f1=0.9853768278965129,
best_f1=0.9854096520763187
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.
```

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

```
2473it [00:14, 165.89it/s]
```

```
load_f1 = 0.9842342342342343
```

```
real_f1 = 0.9864864864864865
```

```
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, 187.41it/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: 0.6475526690483093
step: 10, loss: 0.5194084644317627
step: 20, loss: 0.5812554359436035
step: 30, loss: 0.3102135956287384
step: 40, loss: 0.28648480772972107
step: 50, loss: 0.36937415599823
step: 60, loss: 0.13737131655216217
step: 70, loss: 0.10971707850694656
step: 80, loss: 0.1269596666097641
step: 90, loss: 0.3133040964603424
step: 100, loss: 0.046357687562704086
step: 110, loss: 0.12801088392734528
step: 120, loss: 0.13510605692863464
step: 130, loss: 0.05705660209059715
step: 140, loss: 0.048300810158252716
step: 150, loss: 0.10401435196399689
step: 160, loss: 0.04120679199695587
step: 170, loss: 0.26204371452331543
step: 180, loss: 0.07638775557279587
step: 190, loss: 0.03451981395483017
step: 200, loss: 0.14604265987873077
step: 210, loss: 0.07636280357837677
step: 220, loss: 0.2194308489561081
step: 230, loss: 0.17632916569709778
step: 240, loss: 0.06940022855997086
step: 250, loss: 0.045930683612823486
step: 260, loss: 0.14694996178150177
step: 270, loss: 0.028020335361361504
step: 280, loss: 0.047457654029130936
step: 290, loss: 0.2523806095123291
step: 300, loss: 0.11165594309568405

```

```

step: 310, loss: 0.1951006054878235
step: 320, loss: 0.05837387591600418
step: 330, loss: 0.057944945991039276
step: 340, loss: 0.071994848549366
step: 350, loss: 0.055454425513744354
step: 360, loss: 0.10502538830041885
step: 370, loss: 0.12246477603912354
step: 380, loss: 0.009781117551028728
step: 390, loss: 0.1604590117931366
step: 400, loss: 0.31134891510009766
step: 410, loss: 0.08445728570222855
step: 420, loss: 0.06073335185647011
step: 430, loss: 0.19966831803321838
step: 440, loss: 0.02099340222775936
step: 450, loss: 0.03957200050354004
step: 460, loss: 0.10918697714805603
step: 470, loss: 0.14604565501213074
step: 480, loss: 0.09011848270893097
step: 490, loss: 0.239435076713562
step: 500, loss: 0.07203640788793564
step: 510, loss: 0.04118490591645241
step: 520, loss: 0.04691878706216812
step: 530, loss: 0.003694736398756504
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9129411764705881, f1=0.9072941176470588,
best_f1=0.9072941176470588
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.15728911757469177
step: 10, loss: 0.11061270534992218
step: 20, loss: 0.12922589480876923
step: 30, loss: 0.055941883474588394
step: 40, loss: 0.09487546980381012
step: 50, loss: 0.14391423761844635
step: 60, loss: 0.030963772907853127
step: 70, loss: 0.09032680839300156
step: 80, loss: 0.11511018872261047
step: 90, loss: 0.02245618775486946
step: 100, loss: 0.030970066785812378
step: 110, loss: 0.0697912648320198
step: 120, loss: 0.0937146544456482

```

```

step: 130, loss: 0.07676251977682114
step: 140, loss: 0.039088450372219086
step: 150, loss: 0.06688801944255829
step: 160, loss: 0.02440699003636837
step: 170, loss: 0.015143631026148796
step: 180, loss: 0.023586289957165718
step: 190, loss: 0.09224667400121689
step: 200, loss: 0.015205373987555504
step: 210, loss: 0.08214057981967926
step: 220, loss: 0.07118942588567734
step: 230, loss: 0.005542594939470291
step: 240, loss: 0.12378881126642227
step: 250, loss: 0.0404609851539135
step: 260, loss: 0.004401014186441898
step: 270, loss: 0.1385708749294281
step: 280, loss: 0.0224036555737257
step: 290, loss: 0.08603270351886749
step: 300, loss: 0.15425458550453186
step: 310, loss: 0.012561622075736523
step: 320, loss: 0.08103913813829422
step: 330, loss: 0.0716654509305954
step: 340, loss: 0.01530509814620018
step: 350, loss: 0.001814250135794282
step: 360, loss: 0.0280461348593235
step: 370, loss: 0.09600129723548889
step: 380, loss: 0.043592147529125214
step: 390, loss: 0.05747106298804283
step: 400, loss: 0.08506372570991516
step: 410, loss: 0.023884059861302376
step: 420, loss: 0.03390556946396828
step: 430, loss: 0.013480843976140022
step: 440, loss: 0.03056824952363968
step: 450, loss: 0.03936490789055824
step: 460, loss: 0.09872011840343475
step: 470, loss: 0.003796419594436884
step: 480, loss: 0.19550542533397675
step: 490, loss: 0.0235460102558136
step: 500, loss: 0.3559230864048004
step: 510, loss: 0.03170171007514
step: 520, loss: 0.058750394731760025
step: 530, loss: 0.1061643660068512
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.918869644484959, f1=0.9142335766423357,  
best_f1=0.9142335766423357  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 0.19976618885993958  
step: 10, loss: 0.0820949375629425  
step: 20, loss: 0.08761131763458252  
step: 30, loss: 0.1924351453781128  
step: 40, loss: 0.008697974495589733  
step: 50, loss: 0.03972434252500534  
step: 60, loss: 0.014728819951415062  
step: 70, loss: 0.01710415631532669  
step: 80, loss: 0.00793164037168026  
step: 90, loss: 0.005637395661324263  
step: 100, loss: 0.044168781489133835  
step: 110, loss: 0.004270361270755529  
step: 120, loss: 0.007858308963477612  
step: 130, loss: 0.009285910986363888  
step: 140, loss: 0.027796056121587753  
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0  
step: 150, loss: 0.048940662294626236  
step: 160, loss: 0.00344594893977046  
step: 170, loss: 0.07017601281404495  
step: 180, loss: 0.043610554188489914  
step: 190, loss: 0.01241295412182808  
step: 200, loss: 0.11310680955648422  
step: 210, loss: 0.029159672558307648  
step: 220, loss: 0.06354827433824539  
step: 230, loss: 0.08465202152729034  
step: 240, loss: 0.0019313807133585215  
step: 250, loss: 0.02282462641596794  
step: 260, loss: 0.03269397094845772  
step: 270, loss: 0.0036916041281074286  
step: 280, loss: 0.06213121861219406  
step: 290, loss: 0.006771902088075876  
step: 300, loss: 0.028781859204173088  
step: 310, loss: 0.045706938952207565  
step: 320, loss: 0.05036931857466698  
step: 330, loss: 0.01942676678299904  
step: 340, loss: 0.09785372763872147  
step: 350, loss: 0.01760297454893589  
step: 360, loss: 0.05505107343196869  
step: 370, loss: 0.013960686512291431  
step: 380, loss: 0.009744953364133835  
step: 390, loss: 0.01437817607074976  
step: 400, loss: 0.015125283971428871  
step: 410, loss: 0.03425661474466324
```



```

step: 420, loss: 0.12420769780874252
step: 430, loss: 0.050870586186647415
step: 440, loss: 0.019549265503883362
step: 450, loss: 0.09778425097465515
step: 460, loss: 0.03343911096453667
step: 470, loss: 0.16689719259738922
step: 480, loss: 0.1302901953458786
step: 490, loss: 0.04413819685578346
step: 500, loss: 0.024071233347058296
step: 510, loss: 0.0253940187394619
step: 520, loss: 0.013018791563808918
step: 530, loss: 0.11614716798067093
/usr/local/lib/python3.7/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=0.912639405204461, f1=0.9171322160148976,
best_f1=0.9142335766423357
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006903354078531265
step: 10, loss: 0.012147132307291031
step: 20, loss: 0.0021262443624436855
step: 30, loss: 0.0022768815979361534
step: 40, loss: 0.011285101994872093
step: 50, loss: 0.003503953106701374
step: 60, loss: 0.001661324524320662
step: 70, loss: 0.11794069409370422
step: 80, loss: 0.014990943484008312
step: 90, loss: 0.03691495209932327
step: 100, loss: 0.020360112190246582
step: 110, loss: 0.05424991995096207
step: 120, loss: 0.00034418547875247896
step: 130, loss: 0.000971839064732194
step: 140, loss: 0.010690946131944656
step: 150, loss: 0.030806317925453186
step: 160, loss: 0.022411614656448364
step: 170, loss: 0.017796728760004044
step: 180, loss: 0.0004859252367168665
step: 190, loss: 0.05085033178329468
step: 200, loss: 0.0011702632764354348
step: 210, loss: 0.1291157752275467
step: 220, loss: 0.002084763254970312
step: 230, loss: 0.021368341520428658
step: 240, loss: 0.008010433986783028
step: 250, loss: 0.20227251946926117
step: 260, loss: 0.0012724354164674878
step: 270, loss: 0.018110981211066246

```

```

step: 280, loss: 0.0792132019996643
step: 290, loss: 0.0762021392583847
step: 300, loss: 0.0010740881552919745
step: 310, loss: 0.02886449359357357
step: 320, loss: 0.002679713536053896
step: 330, loss: 0.005405691917985678
step: 340, loss: 0.0007742526941001415
step: 350, loss: 0.00915561243891716
step: 360, loss: 0.002197782974690199
step: 370, loss: 0.0039216154254972935
step: 380, loss: 0.0005355955800041556
step: 390, loss: 0.004137495532631874
step: 400, loss: 0.0032799215987324715
step: 410, loss: 0.060222119092941284
step: 420, loss: 0.009503178298473358
step: 430, loss: 0.06631256639957428
step: 440, loss: 0.0012272284366190434
step: 450, loss: 0.0028488347306847572
step: 460, loss: 0.007946287281811237
step: 470, loss: 0.006103556137531996
step: 480, loss: 0.002278259489685297
step: 490, loss: 0.002272023120895028
step: 500, loss: 0.0022970796562731266
step: 510, loss: 0.04699166864156723
step: 520, loss: 0.07630971074104309
step: 530, loss: 0.026548707857728004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9189695550351288, f1=0.9094321914594088,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.034303952008485794
step: 10, loss: 0.031059468165040016
step: 20, loss: 0.013501479290425777
step: 30, loss: 0.0066389041021466255
step: 40, loss: 0.05039110407233238
step: 50, loss: 0.014609509147703648
step: 60, loss: 0.020137421786785126
step: 70, loss: 0.06300407648086548
step: 80, loss: 0.0006804143777117133
step: 90, loss: 0.002005471847951412

```

```
step: 100, loss: 0.012696090154349804
step: 110, loss: 0.0007549701840616763
step: 120, loss: 0.02213624306023121
step: 130, loss: 0.0013526736292988062
step: 140, loss: 0.007536903489381075
step: 150, loss: 0.0012632521102204919
step: 160, loss: 0.001119065098464489
step: 170, loss: 0.007447876036167145
step: 180, loss: 0.0004975746851414442
step: 190, loss: 0.0051772985607385635
step: 200, loss: 0.002542173722758889
step: 210, loss: 0.04376976564526558
step: 220, loss: 0.0010791313834488392
step: 230, loss: 0.009879285469651222
step: 240, loss: 0.031135134398937225
step: 250, loss: 0.006718350574374199
step: 260, loss: 0.022000109776854515
step: 270, loss: 0.0019955073948949575
step: 280, loss: 0.0004760606971103698
step: 290, loss: 0.14881247282028198
step: 300, loss: 0.014348629862070084
step: 310, loss: 0.0036801618989557028
step: 320, loss: 0.019157832488417625
step: 330, loss: 0.09921836107969284
step: 340, loss: 0.002393892500549555
step: 350, loss: 0.0013666331069543958
step: 360, loss: 0.00544959120452404
step: 370, loss: 0.0090364133939147
step: 380, loss: 0.0016467898385599256
step: 390, loss: 0.0007515784818679094
step: 400, loss: 0.0018094347324222326
step: 410, loss: 0.0003310876782052219
step: 420, loss: 0.0013067044783383608
step: 430, loss: 0.0005280300974845886
step: 440, loss: 0.004577052779495716
step: 450, loss: 0.001415921957232058
step: 460, loss: 0.007297406904399395
step: 470, loss: 0.018342524766921997
step: 480, loss: 0.0016014763386920094
step: 490, loss: 0.0009209089330397546
step: 500, loss: 0.010652400553226471
step: 510, loss: 0.030674202367663383
step: 520, loss: 0.0029225987382233143
step: 530, loss: 0.01148254331201315
/usr/local/lib/python3.7/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=0.8976225133430374, f1=0.8847474252084355,
```

```
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00859438069164753
step: 10, loss: 0.00025668839225545526
step: 20, loss: 0.031083494424819946
step: 30, loss: 0.003140011103823781
step: 40, loss: 0.0004734103276859969
step: 50, loss: 0.026746060699224472
step: 60, loss: 0.0008256566361524165
step: 70, loss: 0.0011417287169024348
step: 80, loss: 0.001190517912618816
step: 90, loss: 0.0007684879237785935
step: 100, loss: 0.004496612120419741
step: 110, loss: 0.059142641723155975
step: 120, loss: 0.00034274545032531023
step: 130, loss: 0.0015135827707126737
step: 140, loss: 0.05535205826163292
step: 150, loss: 0.0010386689100414515
step: 160, loss: 0.013195361010730267
step: 170, loss: 0.0015411829808726907
step: 180, loss: 0.00435964111238718
step: 190, loss: 0.003603446064516902
step: 200, loss: 0.000837191881146282
step: 210, loss: 0.02132531814277172
step: 220, loss: 0.0007057043840177357
step: 230, loss: 0.00025132435257546604
step: 240, loss: 0.010174205526709557
step: 250, loss: 0.00011250031820964068
step: 260, loss: 0.00010776556155178696
step: 270, loss: 0.03616037964820862
step: 280, loss: 0.035410430282354355
step: 290, loss: 0.00023417483316734433
step: 300, loss: 0.00038173282518982887
step: 310, loss: 0.03294592350721359
step: 320, loss: 0.00017396874318365008
step: 330, loss: 0.000299056846415624
step: 340, loss: 0.1343594491481781
step: 350, loss: 0.015233081765472889
step: 360, loss: 0.05749950557947159
step: 370, loss: 0.0016004249919205904
step: 380, loss: 0.0021673724986612797
step: 390, loss: 0.011942695826292038
step: 400, loss: 0.004116641357541084
step: 410, loss: 0.0012052905512973666
step: 420, loss: 0.0006830890197306871
step: 430, loss: 0.0006144685903564095
```

```

step: 440, loss: 6.882245361339301e-05
step: 450, loss: 0.0005076533416286111
step: 460, loss: 0.00029888463905081153
step: 470, loss: 0.0032752761617302895
step: 480, loss: 0.03372107818722725
step: 490, loss: 0.005223463289439678
step: 500, loss: 0.041612379252910614
step: 510, loss: 0.006309003569185734
step: 520, loss: 0.004039199557155371
step: 530, loss: 0.0631539449095726
/usr/local/lib/python3.7/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=0.9168224299065421, f1=0.914365933551708,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00021138740703463554
step: 10, loss: 0.030865512788295746
step: 20, loss: 0.009038502350449562
step: 30, loss: 0.0028522962238639593
step: 40, loss: 0.0021280935034155846
step: 50, loss: 0.0022603515535593033
step: 60, loss: 0.0012286623241379857
step: 70, loss: 0.0012262089876458049
step: 80, loss: 0.000744001823477447
step: 90, loss: 0.00022187686408869922
step: 100, loss: 0.0007984760450199246
step: 110, loss: 0.007247973699122667
step: 120, loss: 0.00019305842579342425
step: 130, loss: 0.0011224589543417096
step: 140, loss: 0.0007738239946775138
step: 150, loss: 0.03312620148062706
step: 160, loss: 0.0005327602848410606
step: 170, loss: 0.009529408067464828
step: 180, loss: 0.013860146515071392
step: 190, loss: 0.005603880155831575
step: 200, loss: 0.000738787348382175
step: 210, loss: 0.04532988741993904
step: 220, loss: 0.0003861208679154515
step: 230, loss: 0.0001624583383090794
step: 240, loss: 0.000511294521857053
step: 250, loss: 0.01257243100553751
step: 260, loss: 0.0004226786259096116
step: 270, loss: 0.008291563019156456
step: 280, loss: 0.000254934246186167
step: 290, loss: 0.00011912180343642831

```

```

step: 300, loss: 0.00045018200762569904
step: 310, loss: 0.0005312298890203238
step: 320, loss: 0.0008488434250466526
step: 330, loss: 0.0009312854381278157
step: 340, loss: 0.01944854110479355
step: 350, loss: 0.0013798049185425043
step: 360, loss: 0.0005721324705518782
step: 370, loss: 0.00020718068117275834
step: 380, loss: 0.01058126986026764
step: 390, loss: 0.0009710180456750095
step: 400, loss: 0.0005619517178274691
step: 410, loss: 0.003171006916090846
step: 420, loss: 0.0010704899905249476
step: 430, loss: 0.00038121023681014776
step: 440, loss: 0.00014395762991625816
step: 450, loss: 0.0001264369348064065
step: 460, loss: 0.00026405329117551446
step: 470, loss: 0.11190611124038696
step: 480, loss: 0.32346418499946594
step: 490, loss: 0.00036436025402508676
step: 500, loss: 0.0003204247332178056
step: 510, loss: 0.0006574979051947594
step: 520, loss: 0.005574389826506376
step: 530, loss: 0.0056871334090828896
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
epoch 7: dev_f1=0.9085027726432533, f1=0.914757103574702,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008161892183125019
step: 10, loss: 0.0021981035824865103
step: 20, loss: 0.00013390230014920235
step: 30, loss: 0.003617379814386368
step: 40, loss: 6.376604869728908e-05
step: 50, loss: 0.007060988340526819
step: 60, loss: 0.01484239473938942
step: 70, loss: 0.00016919299378059804
step: 80, loss: 0.009913679212331772
step: 90, loss: 0.00020645074255298823
step: 100, loss: 4.952052404405549e-05
step: 110, loss: 6.47742926958017e-05
step: 120, loss: 0.0001111838428187184
step: 130, loss: 4.0816452383296564e-05
step: 140, loss: 0.0007649147883057594

```

```
step: 150, loss: 0.00026272441027686
step: 160, loss: 0.0005826317938044667
step: 170, loss: 0.006108174566179514
step: 180, loss: 0.00013319934078026563
step: 190, loss: 0.005233055911958218
step: 200, loss: 0.0027583600021898746
step: 210, loss: 0.05165165662765503
step: 220, loss: 0.003255230840295553
step: 230, loss: 0.0007225411827675998
step: 240, loss: 9.11590686882846e-05
step: 250, loss: 0.00014980587002355605
step: 260, loss: 0.0005630443920381367
step: 270, loss: 0.0004092224990017712
step: 280, loss: 0.04261539503931999
step: 290, loss: 4.8462847189512104e-05
step: 300, loss: 0.06705867499113083
step: 310, loss: 0.04321665316820145
step: 320, loss: 0.0025792468804866076
step: 330, loss: 0.000397457362851128
step: 340, loss: 0.0009953195694833994
step: 350, loss: 0.001012341002933681
step: 360, loss: 6.930910603841767e-05
step: 370, loss: 0.009330583736300468
step: 380, loss: 0.0005323751829564571
step: 390, loss: 0.09089380502700806
step: 400, loss: 0.0003556983429007232
step: 410, loss: 0.00010271271457895637
step: 420, loss: 0.002824322087690234
step: 430, loss: 0.0003479104198049754
step: 440, loss: 0.00011791344149969518
step: 450, loss: 0.00014205207116901875
step: 460, loss: 6.203891098266467e-05
step: 470, loss: 0.00016395142301917076
step: 480, loss: 4.987855572835542e-05
step: 490, loss: 0.01634739153087139
step: 500, loss: 0.0001761188468663022
step: 510, loss: 0.0001671329519012943
step: 520, loss: 0.0016035328153520823
step: 530, loss: 4.0119011828210205e-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=0.9079869219990658, f1=0.9154600653900047,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004371672694105655
```

step: 10, loss: 7.417505548801273e-05
step: 20, loss: 0.0006395485252141953
step: 30, loss: 0.0009607246611267328
step: 40, loss: 0.007309927139431238
step: 50, loss: 8.383204840356484e-05
step: 60, loss: 6.706149724777788e-05
step: 70, loss: 0.0004886408569291234
step: 80, loss: 0.042594168335199356
step: 90, loss: 0.0003232956805732101
step: 100, loss: 4.544291732599959e-05
step: 110, loss: 0.0013132726307958364
step: 120, loss: 0.0007336000562645495
step: 130, loss: 0.000475927023217082
step: 140, loss: 0.008268343284726143
step: 150, loss: 0.0013823435874655843
step: 160, loss: 0.0006200443604029715
step: 170, loss: 0.05145777016878128
step: 180, loss: 0.00011899933451786637
step: 190, loss: 6.754262722097337e-05
step: 200, loss: 0.000317490310408175
step: 210, loss: 6.595835293410346e-05
step: 220, loss: 0.00048510098713450134
step: 230, loss: 0.0003027700586244464
step: 240, loss: 0.0020124625880271196
step: 250, loss: 0.0004322564054746181
step: 260, loss: 0.0030160201713442802
step: 270, loss: 0.0002445354766678065
step: 280, loss: 0.0030341814272105694
step: 290, loss: 0.0024900685530155897
step: 300, loss: 0.00026158447144553065
step: 310, loss: 0.00625130208209157
step: 320, loss: 0.0011897783260792494
step: 330, loss: 0.00014676763385068625
step: 340, loss: 3.15670549753122e-05
step: 350, loss: 0.0009212181903421879
step: 360, loss: 2.550260251155123e-05
step: 370, loss: 0.0018625385127961636
step: 380, loss: 4.490962965064682e-05
step: 390, loss: 6.171272980282083e-05
step: 400, loss: 0.0023990727495402098
step: 410, loss: 3.303799530840479e-05
step: 420, loss: 3.260256926296279e-05
step: 430, loss: 5.1700477342819795e-05
step: 440, loss: 0.0005697357119061053
step: 450, loss: 3.293026384199038e-05
step: 460, loss: 0.00013391293759923428
step: 470, loss: 4.238479596097022e-05
step: 480, loss: 3.1793802918400615e-05


```

step: 490, loss: 3.2094645575853065e-05
step: 500, loss: 0.0016696015372872353
step: 510, loss: 0.00042989078792743385
step: 520, loss: 0.0004452371213119477
step: 530, loss: 5.539916310226545e-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=0.9181094992980814, f1=0.911970190964136,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006327387527562678
step: 10, loss: 0.002874536206945777
step: 20, loss: 0.00015750338206999004
step: 30, loss: 5.089142723591067e-05
step: 40, loss: 0.01245136745274067
step: 50, loss: 7.135529449442402e-05
step: 60, loss: 2.8843987820437178e-05
step: 70, loss: 4.4989857997279614e-05
step: 80, loss: 0.00015731978055555373
step: 90, loss: 0.002465380821377039
step: 100, loss: 3.591025961213745e-05
step: 110, loss: 3.624162127380259e-05
step: 120, loss: 6.012593803461641e-05
step: 130, loss: 3.1886971555650234e-05
step: 140, loss: 0.17289046943187714
step: 150, loss: 0.00010355298581998795
step: 160, loss: 0.0003673071041703224
step: 170, loss: 0.00023711164249107242
step: 180, loss: 0.001549050910398364
step: 190, loss: 0.000267320399871096
step: 200, loss: 0.010186060331761837
step: 210, loss: 0.00010821336036315188
step: 220, loss: 0.00028093188302591443
step: 230, loss: 0.001608367427252233
step: 240, loss: 9.616761235520244e-05
step: 250, loss: 3.724326234078035e-05
step: 260, loss: 0.00023918182705529034
step: 270, loss: 0.00035474926698952913
step: 280, loss: 0.0010498351184651256
step: 290, loss: 3.141437628073618e-05
step: 300, loss: 0.0003781050618272275
step: 310, loss: 4.355792043497786e-05
step: 320, loss: 0.010538417845964432
step: 330, loss: 2.4768929506535642e-05
step: 340, loss: 8.455736679024994e-05

```

```

step: 350, loss: 9.659736679168418e-05
step: 360, loss: 0.00020756985759362578
step: 370, loss: 0.0009468315402045846
step: 380, loss: 0.0010674792574718595
step: 390, loss: 0.024636121466755867
step: 400, loss: 0.001711286255158484
step: 410, loss: 0.0007059540948830545
step: 420, loss: 0.0098556624725461
step: 430, loss: 5.4747684771427885e-05
step: 440, loss: 0.004472455009818077
step: 450, loss: 0.00013354394468478858
step: 460, loss: 0.00017371762078255415
step: 470, loss: 7.426935917465016e-05
step: 480, loss: 4.8711361159803346e-05
step: 490, loss: 8.08742770459503e-05
step: 500, loss: 0.0024531958624720573
step: 510, loss: 5.4534692026209086e-05
step: 520, loss: 8.727952081244439e-05
step: 530, loss: 0.00039252961869351566
/usr/local/lib/python3.7/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=0.9113323850165956, f1=0.9101283880171184,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0026120280381292105
step: 10, loss: 6.169656262500212e-05
step: 20, loss: 4.4379641622072086e-05
step: 30, loss: 3.2474883482791483e-05
step: 40, loss: 0.0017637934070080519
step: 50, loss: 0.00010025978554040194
step: 60, loss: 0.008205479942262173
step: 70, loss: 2.7238464099355042e-05
step: 80, loss: 2.3025499103823677e-05
step: 90, loss: 0.002746188547462225
step: 100, loss: 0.00043296627700328827
step: 110, loss: 5.162119850865565e-05
step: 120, loss: 0.00012714246986433864
step: 130, loss: 0.0001328450598521158
step: 140, loss: 0.04115237668156624
step: 150, loss: 2.6526344299782068e-05
step: 160, loss: 3.783759893849492e-05
step: 170, loss: 9.012466762214899e-05
step: 180, loss: 0.0003535900905262679
step: 190, loss: 0.00013664495781995356
step: 200, loss: 0.02363531105220318

```

```

step: 210, loss: 0.00017071266483981162
step: 220, loss: 0.0041383784264326096
step: 230, loss: 9.06982459127903e-05
step: 240, loss: 0.00026738987071439624
step: 250, loss: 3.4144159144489095e-05
step: 260, loss: 4.540822919807397e-05
step: 270, loss: 0.0001253620139323175
step: 280, loss: 9.040463191922754e-05
step: 290, loss: 0.00012706495181191713
step: 300, loss: 0.000779221358243376
step: 310, loss: 5.803669409942813e-05
step: 320, loss: 0.00053317507263273
step: 330, loss: 0.00016793468967080116
step: 340, loss: 0.0008949824841693044
step: 350, loss: 0.0008875555940903723
step: 360, loss: 0.0006183020304888487
step: 370, loss: 0.06932580471038818
step: 380, loss: 0.00015069539949763566
step: 390, loss: 8.285077637992799e-05
step: 400, loss: 0.00021565098722930998
step: 410, loss: 9.042817691806704e-05
step: 420, loss: 0.004548641387373209
step: 430, loss: 0.0018959176959469914
step: 440, loss: 0.0002067666791845113
step: 450, loss: 0.0129261938855052
step: 460, loss: 0.00424047838896513
step: 470, loss: 4.617167360265739e-05
step: 480, loss: 0.0004593001212924719
step: 490, loss: 0.0023483210243284702
step: 500, loss: 0.00018102661124430597
step: 510, loss: 0.0001581309479661286
step: 520, loss: 6.30921422271058e-05
step: 530, loss: 3.344021024531685e-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=0.9156626506024096, f1=0.9152073732718895,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006200377247296274
step: 10, loss: 2.8146128897788003e-05
step: 20, loss: 0.00017681684403214604
step: 30, loss: 5.541515565710142e-05
step: 40, loss: 3.739957173820585e-05
step: 50, loss: 0.0015049951616674662
step: 60, loss: 9.405118908034638e-05

```

step: 70, loss: 3.6763158277608454e-05
step: 80, loss: 3.7794627132825553e-05
step: 90, loss: 5.3987438150215894e-05
step: 100, loss: 7.524218381149694e-05
step: 110, loss: 3.505633139866404e-05
step: 120, loss: 0.0007504229433834553
step: 130, loss: 0.001396653475239873
step: 140, loss: 2.0391882571857423e-05
step: 150, loss: 1.6577303540543653e-05
step: 160, loss: 1.6633188351988792e-05
step: 170, loss: 2.7506677724886686e-05
step: 180, loss: 3.0973289540270343e-05
step: 190, loss: 0.006682906299829483
step: 200, loss: 0.0003531241382006556
step: 210, loss: 5.0634283979889005e-05
step: 220, loss: 0.00012844018056057394
step: 230, loss: 6.890261283842847e-05
step: 240, loss: 0.000489668978843838
step: 250, loss: 1.9821891328319907e-05
step: 260, loss: 1.7866232155938633e-05
step: 270, loss: 0.0001494480384280905
step: 280, loss: 9.416676039109007e-05
step: 290, loss: 2.9141259801690467e-05
step: 300, loss: 0.00024460643180646
step: 310, loss: 6.855295941932127e-05
step: 320, loss: 2.8643562473007478e-05
step: 330, loss: 2.4265771571663208e-05
step: 340, loss: 2.074938856821973e-05
step: 350, loss: 0.00011154602543683723
step: 360, loss: 0.012027988210320473
step: 370, loss: 0.0006641854415647686
step: 380, loss: 2.5201014068443328e-05
step: 390, loss: 1.8868251572712325e-05
step: 400, loss: 0.00010929592099273577
step: 410, loss: 0.00017020072846207768
step: 420, loss: 3.457842831267044e-05
step: 430, loss: 8.063685527304187e-05
step: 440, loss: 1.7814056263887323e-05
step: 450, loss: 4.211467603454366e-05
step: 460, loss: 2.0827550542890094e-05
step: 470, loss: 2.2473899662145413e-05
step: 480, loss: 0.032257188111543655
step: 490, loss: 6.809565093135461e-05
step: 500, loss: 2.242554910480976e-05
step: 510, loss: 0.014646990224719048
step: 520, loss: 2.5845298296189867e-05
step: 530, loss: 0.00035540247336030006
/usr/local/lib/python3.7/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=0.9183577159037282, f1=0.9135338345864661,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038324533961713314
step: 10, loss: 0.0001364270137855783
step: 20, loss: 1.9758233975153416e-05
step: 30, loss: 3.786891829804517e-05
step: 40, loss: 2.143109122698661e-05
step: 50, loss: 2.5398247089469805e-05
step: 60, loss: 0.0008957526879385114
step: 70, loss: 2.9623968657688238e-05
step: 80, loss: 0.00025584158720448613
step: 90, loss: 0.00047344379709102213
step: 100, loss: 2.8325908715487458e-05
step: 110, loss: 1.516170050308574e-05
step: 120, loss: 2.02130431716796e-05
step: 130, loss: 1.6316464098053984e-05
step: 140, loss: 0.00021948548965156078
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.06960725039243698
step: 160, loss: 0.00015390230691991746
step: 170, loss: 4.552517566480674e-05
step: 180, loss: 8.855920168571174e-05
step: 190, loss: 8.76143021741882e-05
step: 200, loss: 3.709373777383007e-05
step: 210, loss: 6.180688069434837e-05
step: 220, loss: 1.8383885617367923e-05
step: 230, loss: 6.656058394582942e-05
step: 240, loss: 0.00029027790878899395
step: 250, loss: 1.3511502402252518e-05
step: 260, loss: 1.9084060113527812e-05
step: 270, loss: 1.6003607015591115e-05
step: 280, loss: 3.371827187947929e-05
step: 290, loss: 0.00018226777319796383
step: 300, loss: 4.2954652599291876e-05
step: 310, loss: 0.0005916185327805579
step: 320, loss: 0.0008171998779289424
step: 330, loss: 0.002351078437641263
step: 340, loss: 4.320326479501091e-05
step: 350, loss: 3.502541221678257e-05
step: 360, loss: 2.794940519379452e-05
step: 370, loss: 5.355244502425194e-05
step: 380, loss: 0.00016908282123040408
step: 390, loss: 0.00023334268189501017

```

```

step: 400, loss: 1.7996288079302758e-05
step: 410, loss: 6.338913954095915e-05
step: 420, loss: 2.2980386347626336e-05
step: 430, loss: 7.223717693705112e-05
step: 440, loss: 0.003919374197721481
step: 450, loss: 6.962820771150291e-05
step: 460, loss: 0.0002231677935924381
step: 470, loss: 0.0009941416792571545
step: 480, loss: 2.8591601221705787e-05
step: 490, loss: 9.571696864441037e-05
step: 500, loss: 1.3112880878907163e-05
step: 510, loss: 5.9667941968655214e-05
step: 520, loss: 0.0010325354523956776
step: 530, loss: 1.4230446140572894e-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=0.9153846153846154, f1=0.910569105691057,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.899928409140557e-05
step: 10, loss: 1.8066806660499424e-05
step: 20, loss: 0.0009561721817590296
step: 30, loss: 4.52007916464936e-05
step: 40, loss: 0.00031165804830379784
step: 50, loss: 7.018347241682932e-05
step: 60, loss: 2.1836551240994595e-05
step: 70, loss: 4.7926572733558714e-05
step: 80, loss: 3.098417437286116e-05
step: 90, loss: 3.6090110370423645e-05
step: 100, loss: 6.56280608382076e-05
step: 110, loss: 2.327433139726054e-05
step: 120, loss: 0.002799044596031308
step: 130, loss: 5.6057226174743846e-05
step: 140, loss: 1.5236194485623855e-05
step: 150, loss: 1.6390788005082868e-05
step: 160, loss: 2.779243004624732e-05
step: 170, loss: 3.4858585422625765e-05
step: 180, loss: 0.0004929906572215259
step: 190, loss: 3.3584681659704074e-05
step: 200, loss: 2.8560307328007184e-05
step: 210, loss: 2.9374696168815717e-05
step: 220, loss: 0.00036541357985697687
step: 230, loss: 2.6072568289237097e-05
step: 240, loss: 2.3401182261295617e-05
step: 250, loss: 5.280922050587833e-05

```

```

step: 260, loss: 0.0039491960778832436
step: 270, loss: 3.153272336930968e-05
step: 280, loss: 2.973186201415956e-05
step: 290, loss: 6.853900413261726e-05
step: 300, loss: 4.92453618790023e-05
step: 310, loss: 2.381138619966805e-05
step: 320, loss: 6.212318839970976e-05
step: 330, loss: 3.471978197921999e-05
step: 340, loss: 1.9233337297919206e-05
step: 350, loss: 4.4660693674813956e-05
step: 360, loss: 0.00010850441321963444
step: 370, loss: 7.308884960366413e-05
step: 380, loss: 1.7858739738585427e-05
step: 390, loss: 0.013260887004435062
step: 400, loss: 2.9879058274673298e-05
step: 410, loss: 3.111628393526189e-05
step: 420, loss: 8.56599654071033e-05
step: 430, loss: 2.5949557311832905e-05
step: 440, loss: 2.740605123108253e-05
step: 450, loss: 0.000202861032448709
step: 460, loss: 4.21556833316572e-05
step: 470, loss: 3.8236266846070066e-05
step: 480, loss: 3.0630930268671364e-05
step: 490, loss: 1.9024615539819933e-05
step: 500, loss: 1.3425776160147507e-05
step: 510, loss: 5.657122892444022e-05
step: 520, loss: 1.5202686881821137e-05
step: 530, loss: 0.00012323373812250793
/usr/local/lib/python3.7/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=0.9173435260391782, f1=0.9158345221112696,
best_f1=0.9094321914594088
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.5665711291367188e-05
step: 10, loss: 0.0006758849485777318
step: 20, loss: 1.676722058618907e-05
step: 30, loss: 3.122552880086005e-05
step: 40, loss: 0.06981639564037323
step: 50, loss: 0.0002254262799397111
step: 60, loss: 1.2259804861969315e-05
step: 70, loss: 0.0002021319087361917
step: 80, loss: 8.443683327641338e-05
step: 90, loss: 1.782518666004762e-05
step: 100, loss: 3.1811017834115773e-05
step: 110, loss: 1.9233260900364257e-05

```

```
step: 120, loss: 1.5832289136596955e-05
step: 130, loss: 0.0013349582441151142
step: 140, loss: 9.53461931203492e-05
step: 150, loss: 2.1695654140785336e-05
step: 160, loss: 2.5319523047073744e-05
step: 170, loss: 2.1884854504605755e-05
step: 180, loss: 1.5824753063498065e-05
step: 190, loss: 2.5647766960901208e-05
step: 200, loss: 2.728707841015421e-05
step: 210, loss: 2.427580693620257e-05
step: 220, loss: 1.4286307305155788e-05
step: 230, loss: 1.8093462131218985e-05
step: 240, loss: 1.6480425983900204e-05
step: 250, loss: 1.629416692594532e-05
step: 260, loss: 1.6104193491628394e-05
step: 270, loss: 1.4629049474024214e-05
step: 280, loss: 4.002493369625881e-05
step: 290, loss: 1.4178270248521585e-05
step: 300, loss: 1.8171549527323805e-05
step: 310, loss: 0.001645322423428297
step: 320, loss: 1.4081325389270205e-05
step: 330, loss: 0.00015273746976163238
step: 340, loss: 1.455452547816094e-05
step: 350, loss: 1.168614107882604e-05
step: 360, loss: 0.00020324117213021964
step: 370, loss: 4.6332737838383764e-05
step: 380, loss: 7.849063695175573e-05
step: 390, loss: 0.0001287843333557248
step: 400, loss: 2.1706673578592017e-05
step: 410, loss: 6.514896813314408e-05
step: 420, loss: 3.988268872490153e-05
step: 430, loss: 2.8702186682494357e-05
step: 440, loss: 2.7275185857433826e-05
step: 450, loss: 1.3332590242498554e-05
step: 460, loss: 1.6569774743402377e-05
step: 470, loss: 0.004403131082653999
step: 480, loss: 1.1540841114765499e-05
step: 490, loss: 1.6282991055049933e-05
step: 500, loss: 2.8075630325474776e-05
step: 510, loss: 1.913267988129519e-05
step: 520, loss: 0.0001167122318292968
step: 530, loss: 1.8965100025525317e-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=0.9184060721062618, f1=0.9217719132893496,
best_f1=0.9094321914594088
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, 193.73it/s]
load_f1 = 0.9137055837563451
real_f1 = 0.9118597138901707
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, 191.11it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.5546267628669739
step: 10, loss: 0.3706984221935272
step: 20, loss: 0.3917650580406189
step: 30, loss: 0.3281698226928711
step: 40, loss: 0.19150260090827942
step: 50, loss: 0.46106916666030884
step: 60, loss: 0.3170819878578186
step: 70, loss: 0.20616382360458374
step: 80, loss: 0.19859230518341064
step: 90, loss: 0.41078051924705505
step: 100, loss: 0.37071073055267334
step: 110, loss: 0.28437939286231995
step: 120, loss: 0.24777436256408691
step: 130, loss: 0.3017869293689728
step: 140, loss: 0.21023693680763245
step: 150, loss: 0.28309738636016846
step: 160, loss: 0.24851208925247192
step: 170, loss: 0.28546035289764404
step: 180, loss: 0.11384940892457962
step: 190, loss: 0.22546467185020447
step: 200, loss: 0.17514555156230927
step: 210, loss: 0.26512837409973145
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5330739299610895, f1=0.5402298850574712,
best_f1=0.5402298850574712
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.13939817249774933
step: 10, loss: 0.2532702386379242
step: 20, loss: 0.19068442285060883
step: 30, loss: 0.2951449155807495
step: 40, loss: 0.15639862418174744
step: 50, loss: 0.10967954993247986
step: 60, loss: 0.3411538600921631
step: 70, loss: 0.08769296109676361
step: 80, loss: 0.14382494986057281
step: 90, loss: 0.17843279242515564
step: 100, loss: 0.060488175600767136
step: 110, loss: 0.09324368834495544
step: 120, loss: 0.16769948601722717
step: 130, loss: 0.06563220918178558
step: 140, loss: 0.21980564296245575
step: 150, loss: 0.29852813482284546
step: 160, loss: 0.24032190442085266
step: 170, loss: 0.18672168254852295
step: 180, loss: 0.17469710111618042
step: 190, loss: 0.22362984716892242
step: 200, loss: 0.06321366131305695
step: 210, loss: 0.1308053582906723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5557809330628803, f1=0.5269978401727863,
best_f1=0.5269978401727863
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08784407377243042
step: 10, loss: 0.14606241881847382
step: 20, loss: 0.10036443173885345
step: 30, loss: 0.19921965897083282
step: 40, loss: 0.05107968673110008
step: 50, loss: 0.06091108173131943
step: 60, loss: 0.35216930508613586
step: 70, loss: 0.1828518658876419
step: 80, loss: 0.12617270648479462
step: 90, loss: 0.06574933230876923
step: 100, loss: 0.1718505173921585
step: 110, loss: 0.16171517968177795
step: 120, loss: 0.22614246606826782
step: 130, loss: 0.194395050406456

```

```

step: 140, loss: 0.09833331406116486
step: 150, loss: 0.2665715217590332
step: 160, loss: 0.037360768765211105
step: 170, loss: 0.22010372579097748
step: 180, loss: 0.14619113504886627
step: 190, loss: 0.13746747374534607
step: 200, loss: 0.041152313351631165
step: 210, loss: 0.14536090195178986
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5631469979296067, f1=0.559322033898305,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2135348916053772
step: 10, loss: 0.10975370556116104
step: 20, loss: 0.2515109181404114
step: 30, loss: 0.24461764097213745
step: 40, loss: 0.06556086242198944
step: 50, loss: 0.10207764804363251
step: 60, loss: 0.16753526031970978
step: 70, loss: 0.23647283017635345
step: 80, loss: 0.11663813143968582
step: 90, loss: 0.04609997570514679
step: 100, loss: 0.21941693127155304
step: 110, loss: 0.07752154767513275
step: 120, loss: 0.09510241448879242
step: 130, loss: 0.21711428463459015
step: 140, loss: 0.1296544075012207
step: 150, loss: 0.1892731636762619
step: 160, loss: 0.034320633858442307
step: 170, loss: 0.159883052110672
step: 180, loss: 0.36177924275398254
step: 190, loss: 0.1853923350572586
step: 200, loss: 0.22971196472644806
step: 210, loss: 0.1961688995361328
/usr/local/lib/python3.7/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=0.5581395348837208, f1=0.5434782608695652,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07986637949943542
step: 10, loss: 0.012438011355698109
step: 20, loss: 0.08920589089393616
step: 30, loss: 0.028478896245360374
step: 40, loss: 0.1120123490691185
step: 50, loss: 0.017898626625537872
step: 60, loss: 0.1224168911576271
step: 70, loss: 0.22434523701667786
step: 80, loss: 0.23916520178318024
step: 90, loss: 0.1250927746295929
step: 100, loss: 0.037918638437986374
step: 110, loss: 0.12032382935285568
step: 120, loss: 0.05205977335572243
step: 130, loss: 0.09760177880525589
step: 140, loss: 0.08427602052688599
step: 150, loss: 0.022042715921998024
step: 160, loss: 0.11283376067876816
step: 170, loss: 0.08877435326576233
step: 180, loss: 0.10119622200727463
step: 190, loss: 0.08203335851430893
step: 200, loss: 0.06177103519439697
step: 210, loss: 0.04315566644072533
/usr/local/lib/python3.7/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=0.5410821643286573, f1=0.5203252032520326,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.048537299036979675
step: 10, loss: 0.0683123990893364
step: 20, loss: 0.04246288165450096
step: 30, loss: 0.012206786312162876
step: 40, loss: 0.06857240945100784
step: 50, loss: 0.010477036237716675
step: 60, loss: 0.05532844364643097
step: 70, loss: 0.0021741942036896944
step: 80, loss: 0.09084373712539673
step: 90, loss: 0.3633959889411926
step: 100, loss: 0.024137156084179878
step: 110, loss: 0.012963572517037392
step: 120, loss: 0.07262597978115082
step: 130, loss: 0.09738346934318542
step: 140, loss: 0.16546130180358887
step: 150, loss: 0.1381332278251648

```

```

step: 160, loss: 0.028539394959807396
step: 170, loss: 0.0986403077840805
step: 180, loss: 0.0051866560243070126
step: 190, loss: 0.1396821290254593
step: 200, loss: 0.009322868660092354
step: 210, loss: 0.1412615180015564
/usr/local/lib/python3.7/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=0.5274261603375527, f1=0.49783549783549785,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01170356199145317
step: 10, loss: 0.006382729858160019
step: 20, loss: 0.011990259401500225
step: 30, loss: 0.021737249568104744
step: 40, loss: 0.005095252767205238
step: 50, loss: 0.10858387500047684
step: 60, loss: 0.03763938322663307
step: 70, loss: 0.014209376648068428
step: 80, loss: 0.07067367434501648
step: 90, loss: 0.022653963416814804
step: 100, loss: 0.0008143142331391573
step: 110, loss: 0.006224920507520437
step: 120, loss: 0.08020038157701492
step: 130, loss: 0.09044057130813599
step: 140, loss: 0.03073640540242195
step: 150, loss: 0.04851997643709183
step: 160, loss: 0.02267376147210598
step: 170, loss: 0.06348323076963425
step: 180, loss: 0.01173819787800312
step: 190, loss: 0.029191166162490845
step: 200, loss: 0.01724427565932274
step: 210, loss: 0.11271966248750687
/usr/local/lib/python3.7/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=0.5234042553191489, f1=0.502145922746781,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03145546838641167
step: 10, loss: 0.021190855652093887
step: 20, loss: 0.0019700266420841217
step: 30, loss: 0.034458719193935394

```

```

step: 40, loss: 0.010210487991571426
step: 50, loss: 0.006376994773745537
step: 60, loss: 0.069615438580513
step: 70, loss: 0.19330938160419464
step: 80, loss: 0.02241945080459118
step: 90, loss: 0.0030247357208281755
step: 100, loss: 0.008069063536822796
step: 110, loss: 0.009739420376718044
step: 120, loss: 0.000713236047886312
step: 130, loss: 0.004523781593888998
step: 140, loss: 0.0049616750329732895
step: 150, loss: 0.021038463339209557
step: 160, loss: 0.14868362247943878
step: 170, loss: 0.0031675726640969515
step: 180, loss: 0.032596688717603683
step: 190, loss: 0.0023323078639805317
step: 200, loss: 0.001055573346093297
step: 210, loss: 0.13613830506801605
/usr/local/lib/python3.7/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=0.5352697095435683, f1=0.5161290322580645,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007909407839179039
step: 10, loss: 0.016054445877671242
step: 20, loss: 0.002083503408357501
step: 30, loss: 0.07630036771297455
step: 40, loss: 0.010893273167312145
step: 50, loss: 0.014157356694340706
step: 60, loss: 0.0648505687713623
step: 70, loss: 0.005184843670576811
step: 80, loss: 0.0026157123502343893
step: 90, loss: 0.0087191853672266
step: 100, loss: 0.02004418894648552
step: 110, loss: 0.03763315826654434
step: 120, loss: 0.0005008678999729455
step: 130, loss: 0.07406828552484512
step: 140, loss: 0.0411229282617569
step: 150, loss: 0.059997525066137314
step: 160, loss: 0.0015506462659686804
step: 170, loss: 0.000982865341939032
step: 180, loss: 0.017264772206544876
step: 190, loss: 0.0028825157787650824
step: 200, loss: 0.0015963707119226456
step: 210, loss: 0.009345428086817265

```



```

/usr/local/lib/python3.7/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=0.5566600397614314, f1=0.4861660079051383,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.023831842467188835
step: 10, loss: 0.039551492780447006
step: 20, loss: 0.02853352390229702
step: 30, loss: 0.05760025233030319
step: 40, loss: 0.006763600278645754
step: 50, loss: 0.0006401030113920569
step: 60, loss: 0.012040555477142334
step: 70, loss: 0.013256575912237167
step: 80, loss: 0.003992289770394564
step: 90, loss: 0.0325687937438488
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 100, loss: 0.027860965579748154
step: 110, loss: 0.040174245834350586
step: 120, loss: 0.002053368603810668
step: 130, loss: 0.162983700633049
step: 140, loss: 0.008081983774900436
step: 150, loss: 0.002030531642958522
step: 160, loss: 0.0032904702238738537
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.15426577627658844
step: 180, loss: 0.00406343350186944
step: 190, loss: 0.06998149305582047
step: 200, loss: 0.0008490580949001014
step: 210, loss: 0.03595159947872162
/usr/local/lib/python3.7/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=0.5327313769751693, f1=0.4658823529411765,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03678065165877342
step: 10, loss: 0.018562588840723038
step: 20, loss: 0.030542364344000816
step: 30, loss: 0.00016218902601394802
step: 40, loss: 0.01748044230043888
step: 50, loss: 0.002939770231023431
step: 60, loss: 0.06173865124583244
step: 70, loss: 0.012449365109205246

```

```

step: 80, loss: 0.01901380717754364
step: 90, loss: 0.0005323176155798137
step: 100, loss: 0.001129221636801958
step: 110, loss: 0.008953913114964962
step: 120, loss: 0.028866562992334366
step: 130, loss: 0.011857716366648674
step: 140, loss: 0.00020472059259191155
step: 150, loss: 0.0002160118892788887
step: 160, loss: 0.0003826832980848849
step: 170, loss: 0.10085194557905197
step: 180, loss: 0.02373097464442253
step: 190, loss: 0.005401375703513622
step: 200, loss: 0.0006625454407185316
step: 210, loss: 0.0003190015268046409
/usr/local/lib/python3.7/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=0.5343511450381679, f1=0.4940239043824701,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08060410618782043
step: 10, loss: 0.0007647340535186231
step: 20, loss: 0.00420095631852746
step: 30, loss: 0.003413196187466383
step: 40, loss: 0.01183103397488594
step: 50, loss: 0.0035181534476578236
step: 60, loss: 0.00608406588435173
step: 70, loss: 0.0022330975625663996
step: 80, loss: 0.00026484241243451834
step: 90, loss: 0.00046344890142790973
step: 100, loss: 0.004131681751459837
step: 110, loss: 0.002418381394818425
step: 120, loss: 0.005728491116315126
step: 130, loss: 0.0007227729656733572
step: 140, loss: 0.0003773195203393698
step: 150, loss: 0.013907392509281635
step: 160, loss: 0.00021791439212393016
step: 170, loss: 0.0005826516426168382
step: 180, loss: 0.0002545176539570093
step: 190, loss: 0.003776815254241228
step: 200, loss: 0.0004929135320708156
step: 210, loss: 0.005808765534311533
/usr/local/lib/python3.7/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=0.5514223194748359, f1=0.4801864801864802,

```

```

best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0024107927456498146
step: 10, loss: 0.0003024096367880702
step: 20, loss: 0.024628661572933197
step: 30, loss: 0.010101527906954288
step: 40, loss: 0.0004808926023542881
step: 50, loss: 0.00022008837549947202
step: 60, loss: 0.0017895149067044258
step: 70, loss: 0.028990013524889946
step: 80, loss: 0.0011411180021241307
step: 90, loss: 0.0001292819215450436
step: 100, loss: 0.0010358477011322975
step: 110, loss: 0.0003591643471736461
step: 120, loss: 0.0002536297542974353
step: 130, loss: 0.001593122724443674
step: 140, loss: 0.00025371077936142683
step: 150, loss: 0.017065180465579033
step: 160, loss: 0.007873116061091423
step: 170, loss: 0.057531777769327164
step: 180, loss: 0.018410304561257362
step: 190, loss: 0.0006330916658043861
step: 200, loss: 0.004116215277463198
step: 210, loss: 0.005106380674988031
/usr/local/lib/python3.7/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=0.5415778251599146, f1=0.49777777777777776,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006584248621948063
step: 10, loss: 0.02979096956551075
step: 20, loss: 0.0006333271157927811
step: 30, loss: 0.0013377557042986155
step: 40, loss: 0.00036312316660769284
step: 50, loss: 0.002968903398141265
step: 60, loss: 0.00044109768350608647
step: 70, loss: 0.00021876337996218354
step: 80, loss: 0.0014988427283242345
step: 90, loss: 0.0051051960326731205
step: 100, loss: 0.00022594213078264147
step: 110, loss: 0.0005430894671007991
step: 120, loss: 0.04483213648200035
step: 130, loss: 0.0009842216968536377

```

```

step: 140, loss: 0.0014997866237536073
step: 150, loss: 0.0002354552852921188
step: 160, loss: 0.000824901566375047
step: 170, loss: 0.003196483477950096
step: 180, loss: 0.0024800323881208897
step: 190, loss: 0.00024236507306341082
step: 200, loss: 0.0003370074846316129
step: 210, loss: 0.020334526896476746
/usr/local/lib/python3.7/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=0.5412262156448203, f1=0.5077262693156732,
best_f1=0.559322033898305
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00020486237190198153
step: 10, loss: 0.00011924895807169378
step: 20, loss: 0.0008479770622216165
step: 30, loss: 0.0019613902550190687
step: 40, loss: 0.0004765300254803151
step: 50, loss: 0.0203773844987154
step: 60, loss: 0.01500740461051464
step: 70, loss: 0.0008240934694185853
step: 80, loss: 0.0007144531700760126
step: 90, loss: 0.0007302733138203621
step: 100, loss: 0.015179691836237907
step: 110, loss: 0.00016355101251974702
step: 120, loss: 0.0001900758798001334
step: 130, loss: 0.00018596855807118118
step: 140, loss: 0.00019128907297272235
step: 150, loss: 0.00018121513130608946
step: 160, loss: 0.0006551998085342348
step: 170, loss: 0.000881471554748714
step: 180, loss: 0.0008157175616361201
step: 190, loss: 0.00036491951323114336
step: 200, loss: 0.00039714205195195973
step: 210, loss: 0.02160993032157421
/usr/local/lib/python3.7/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=0.5319148936170213, f1=0.5022421524663678,
best_f1=0.559322033898305
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.")
2293it [00:08, 275.61it/s]
load_f1 = 0.59375
real_f1 = 0.5996055226824457
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, 190.98it/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: 0.5703436136245728
step: 10, loss: 0.36792805790901184
step: 20, loss: 0.32437416911125183
step: 30, loss: 0.4238101840019226
step: 40, loss: 0.4253441393375397
step: 50, loss: 0.30003267526626587
step: 60, loss: 0.2970193028450012
step: 70, loss: 0.2732957899570465
step: 80, loss: 0.2601363956928253
step: 90, loss: 0.302367627620697
step: 100, loss: 0.34254178404808044
step: 110, loss: 0.36359521746635437
step: 120, loss: 0.1917882263660431
step: 130, loss: 0.1880785971879959
step: 140, loss: 0.05620725080370903
step: 150, loss: 0.17305508255958557
step: 160, loss: 0.13374711573123932
step: 170, loss: 0.24767081439495087
step: 180, loss: 0.038666218519210815
step: 190, loss: 0.24717940390110016

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

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

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

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=0.6334310850439884, f1=0.6046511627906976,
best_f1=0.6046511627906976

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

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

step: 0, loss: 0.26283836364746094
step: 10, loss: 0.1301906555891037
step: 20, loss: 0.20547643303871155
step: 30, loss: 0.22031748294830322
step: 40, loss: 0.19080683588981628
step: 50, loss: 0.12322111427783966

```

step: 60, loss: 0.34905603528022766
step: 70, loss: 0.20753420889377594
step: 80, loss: 0.30647528171539307
step: 90, loss: 0.3711014986038208
step: 100, loss: 0.06988214701414108
step: 110, loss: 0.1682816743850708
step: 120, loss: 0.3422984480857849
step: 130, loss: 0.11998416483402252
step: 140, loss: 0.09014297276735306
step: 150, loss: 0.11748626828193665
step: 160, loss: 0.03823975846171379
step: 170, loss: 0.19938983023166656
step: 180, loss: 0.1771990954875946
step: 190, loss: 0.046994224190711975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7183462532299743, f1=0.7076167076167075,
best_f1=0.7076167076167075
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07958737015724182
step: 10, loss: 0.36134809255599976
step: 20, loss: 0.12007302790880203
step: 30, loss: 0.1522369235754013
step: 40, loss: 0.117103710770607
step: 50, loss: 0.20986023545265198
step: 60, loss: 0.09343495219945908
step: 70, loss: 0.09676934033632278
step: 80, loss: 0.11069457232952118
step: 90, loss: 0.08836106956005096
step: 100, loss: 0.13532108068466187
step: 110, loss: 0.04775312542915344
step: 120, loss: 0.032574933022260666
step: 130, loss: 0.008993968367576599
step: 140, loss: 0.016420423984527588
step: 150, loss: 0.17553186416625977
step: 160, loss: 0.12306471168994904
step: 170, loss: 0.13261373341083527
step: 180, loss: 0.05508045852184296
step: 190, loss: 0.06826436519622803
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7341040462427746, f1=0.7521367521367521,
best_f1=0.7521367521367521
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.014259236864745617
step: 10, loss: 0.15118330717086792
step: 20, loss: 0.029089229181408882
step: 30, loss: 0.01315529178828001
step: 40, loss: 0.010238170623779297
step: 50, loss: 0.1584194451570511
step: 60, loss: 0.06854937970638275
step: 70, loss: 0.06270712614059448
step: 80, loss: 0.15260428190231323
step: 90, loss: 0.022347629070281982
step: 100, loss: 0.13577350974082947
step: 110, loss: 0.036925170570611954
step: 120, loss: 0.06890463083982468
step: 130, loss: 0.2423945665359497
step: 140, loss: 0.03347661718726158
step: 150, loss: 0.03694763034582138
step: 160, loss: 0.027798889204859734
step: 170, loss: 0.27024105191230774
step: 180, loss: 0.009282679297029972
step: 190, loss: 0.16927333176136017
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7345844504021448, f1=0.7336683417085428,
best_f1=0.7336683417085428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.089076928794384
step: 10, loss: 0.010937655344605446
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.07044987380504608
step: 30, loss: 0.007144297938793898
step: 40, loss: 0.07719486951828003

```

```

step: 50, loss: 0.18177025020122528
step: 60, loss: 0.12446268647909164
step: 70, loss: 0.0055403090082108974
step: 80, loss: 0.046123385429382324
step: 90, loss: 0.03289204463362694
step: 100, loss: 0.0022686636075377464
step: 110, loss: 0.010759861208498478
step: 120, loss: 0.010490956716239452
step: 130, loss: 0.04443899169564247
step: 140, loss: 0.10976976156234741
step: 150, loss: 0.09923660010099411
step: 160, loss: 0.04803546890616417
step: 170, loss: 0.020180700346827507
step: 180, loss: 0.19711445271968842
step: 190, loss: 0.08608684688806534
/usr/local/lib/python3.7/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=0.7295285359801489, f1=0.7487684729064039,
best_f1=0.7336683417085428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.016036666929721832
step: 10, loss: 0.0464971661567688
step: 20, loss: 0.002707048552110791
step: 30, loss: 0.004113068338483572
step: 40, loss: 0.017644239589571953
step: 50, loss: 0.022174861282110214
step: 60, loss: 0.05365774407982826
step: 70, loss: 0.09488271176815033
step: 80, loss: 0.007040045689791441
step: 90, loss: 0.0668630376458168
step: 100, loss: 0.0020050513558089733
step: 110, loss: 0.12651225924491882
step: 120, loss: 0.06374119967222214
step: 130, loss: 0.014677093364298344
step: 140, loss: 0.01630280539393425
step: 150, loss: 0.0020300892647355795
step: 160, loss: 0.007331767585128546
step: 170, loss: 0.23026441037654877
step: 180, loss: 0.0074007101356983185
step: 190, loss: 0.009552664123475552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7386363636363636, f1=0.7506849315068493,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004324585665017366
step: 10, loss: 0.015859460458159447
step: 20, loss: 0.08962202817201614
step: 30, loss: 0.062330882996320724
step: 40, loss: 0.0106204804033041
step: 50, loss: 0.006024419795721769
step: 60, loss: 0.004823050927370787
step: 70, loss: 0.009285311214625835
step: 80, loss: 0.0300692617893219
step: 90, loss: 0.03772279620170593
step: 100, loss: 0.000827490643132478
step: 110, loss: 0.02337065152823925
step: 120, loss: 0.0058722710236907005
step: 130, loss: 0.0038913448806852102
step: 140, loss: 0.08376806229352951
step: 150, loss: 0.04230375215411186
step: 160, loss: 0.053169459104537964
step: 170, loss: 0.006071748211979866
step: 180, loss: 0.006698683835566044
step: 190, loss: 0.00620218925178051
/usr/local/lib/python3.7/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=0.7044334975369458, f1=0.7383863080684596,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007674987427890301
step: 10, loss: 0.022222299128770828
step: 20, loss: 0.0005535662057809532
step: 30, loss: 0.019369270652532578
step: 40, loss: 0.0009230804280377924
step: 50, loss: 0.00037116589373908937
step: 60, loss: 0.0004694376257248223
step: 70, loss: 0.007240189239382744
step: 80, loss: 0.00012773234630003572
step: 90, loss: 0.0027909378986805677
step: 100, loss: 0.06137893721461296
step: 110, loss: 0.004357069730758667
step: 120, loss: 0.00025914801517501473

```

```

step: 130, loss: 0.0010349731892347336
step: 140, loss: 0.0010446569649502635
step: 150, loss: 0.002445588354021311
step: 160, loss: 0.0030487962067127228
step: 170, loss: 0.03202594816684723
step: 180, loss: 0.03239750862121582
step: 190, loss: 0.07399027049541473
/usr/local/lib/python3.7/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=0.7191601049868767, f1=0.7448979591836734,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03299075737595558
step: 10, loss: 0.010495869442820549
step: 20, loss: 0.0009525843197479844
step: 30, loss: 0.00024332708562724292
step: 40, loss: 0.00025106396060436964
step: 50, loss: 0.004190970212221146
step: 60, loss: 0.032621871680021286
step: 70, loss: 0.00043498011655174196
step: 80, loss: 0.00044501895899884403
step: 90, loss: 0.005987532902508974
step: 100, loss: 0.0024746458511799574
step: 110, loss: 0.059946637600660324
step: 120, loss: 0.059878770262002945
step: 130, loss: 0.0011338673066347837
step: 140, loss: 0.013211927376687527
step: 150, loss: 0.004932434298098087
step: 160, loss: 0.0005081894341856241
step: 170, loss: 0.0011439918307587504
step: 180, loss: 0.0005712947458960116
step: 190, loss: 0.0008832511957734823
/usr/local/lib/python3.7/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=0.7246376811594202, f1=0.7507163323782234,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00026075271307490766
step: 10, loss: 0.0003165117232128978
step: 20, loss: 0.001374400919303298
step: 30, loss: 0.0001425138907507062
step: 40, loss: 0.034811969846487045

```

```

step: 50, loss: 0.007911399938166142
step: 60, loss: 0.00023264317132998258
step: 70, loss: 0.0002465485886204988
step: 80, loss: 0.0007803849293850362
step: 90, loss: 0.00441198842599988
step: 100, loss: 0.0008626877679489553
step: 110, loss: 0.02380412258207798
step: 120, loss: 0.10076631605625153
step: 130, loss: 0.004128111060708761
step: 140, loss: 0.020053386688232422
step: 150, loss: 0.02458023466169834
step: 160, loss: 0.001597483758814633
step: 170, loss: 0.00040693487972021103
step: 180, loss: 0.001430949429050088
step: 190, loss: 0.00020575826056301594
/usr/local/lib/python3.7/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=0.7161803713527852, f1=0.7512953367875648,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0038175852969288826
step: 10, loss: 0.021858714520931244
step: 20, loss: 0.0011790121207013726
step: 30, loss: 0.016415797173976898
step: 40, loss: 0.000193356754607521
step: 50, loss: 0.0019192870240658522
step: 60, loss: 0.0013861296465620399
step: 70, loss: 0.0011478858068585396
step: 80, loss: 0.00018783201812766492
step: 90, loss: 0.0001661512505961582
step: 100, loss: 0.0009411242208443582
step: 110, loss: 0.00042888676398433745
step: 120, loss: 0.00038290268275886774
step: 130, loss: 0.0006470793741755188
step: 140, loss: 0.004053832031786442
step: 150, loss: 0.00011045342398574576
step: 160, loss: 0.002355254255235195
step: 170, loss: 0.00014788127737119794
step: 180, loss: 0.00028515412122942507
step: 190, loss: 0.0005147833726368845
/usr/local/lib/python3.7/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=0.7262569832402234, f1=0.7409470752089136,
best_f1=0.7506849315068493

```

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00018612443818710744
step: 10, loss: 0.009281555190682411
step: 20, loss: 0.00016155745834112167
step: 30, loss: 0.00039626526995562017
step: 40, loss: 0.0007178130908869207
step: 50, loss: 0.0146110150963068
step: 60, loss: 0.0005117481923662126
step: 70, loss: 0.00031762145226821303
step: 80, loss: 0.0002088043256662786
step: 90, loss: 0.00028595083858817816
step: 100, loss: 0.0015736838104203343
step: 110, loss: 0.0007734206155873835
step: 120, loss: 0.00014837317576166242
step: 130, loss: 0.003842931939288974
step: 140, loss: 0.009134369902312756
step: 150, loss: 0.00034207719727419317
step: 160, loss: 0.00014432660827878863
step: 170, loss: 0.21081674098968506
step: 180, loss: 5.221917308517732e-05
step: 190, loss: 0.0005022919503971934
/usr/local/lib/python3.7/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=0.7258064516129032, f1=0.7446808510638299,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005863674450665712
step: 10, loss: 0.0006925062043592334
step: 20, loss: 0.0013302686857059598
step: 30, loss: 0.005415037740021944
step: 40, loss: 0.00014856725465506315
step: 50, loss: 0.001219072612002492
step: 60, loss: 0.00042834412306547165
step: 70, loss: 0.0011701899347826838
step: 80, loss: 0.0003937122237402946
step: 90, loss: 0.0019932540599256754
step: 100, loss: 0.00025744098820723593
step: 110, loss: 0.004965823143720627
step: 120, loss: 0.00331432418897748
step: 130, loss: 0.0020152272190898657
step: 140, loss: 0.0002795204345602542
step: 150, loss: 8.11555510153994e-05
step: 160, loss: 0.00028637261129915714

```

```

step: 170, loss: 0.00025725673185661435
step: 180, loss: 0.000371713686035946
step: 190, loss: 0.0029196874238550663
/usr/local/lib/python3.7/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=0.7075208913649025, f1=0.7150837988826816,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0020050827879458666
step: 10, loss: 0.00017274817219004035
step: 20, loss: 0.0007641944685019553
step: 30, loss: 0.00021661678329110146
step: 40, loss: 0.0004971068119630218
step: 50, loss: 0.001057908870279789
step: 60, loss: 0.001035910565406084
step: 70, loss: 0.00015595744480378926
step: 80, loss: 0.00015039385471027344
step: 90, loss: 0.0032208182383328676
step: 100, loss: 0.06235566735267639
step: 110, loss: 0.0002062873973045498
step: 120, loss: 0.0003070970706176013
step: 130, loss: 0.006311289966106415
step: 140, loss: 0.0012822550488635898
step: 150, loss: 0.00012220583448652178
step: 160, loss: 0.0021614402066916227
step: 170, loss: 0.0003821287536993623
step: 180, loss: 0.0003900978190358728
step: 190, loss: 0.0007516540354117751
/usr/local/lib/python3.7/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=0.7183462532299743, f1=0.739240506329114,
best_f1=0.7506849315068493
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030819906387478113
step: 10, loss: 0.00037800028803758323
step: 20, loss: 0.0017740379553288221
step: 30, loss: 0.0003099735185969621
step: 40, loss: 0.00228230282664299
step: 50, loss: 0.00026161724235862494
step: 60, loss: 0.023381343111395836
step: 70, loss: 0.0019206549040973186
step: 80, loss: 0.0015123924240469933

```

```

step: 90, loss: 0.000549063493963331
step: 100, loss: 0.00042740636854432523
step: 110, loss: 0.0020508256275206804
step: 120, loss: 0.00017718991148285568
step: 130, loss: 0.00015018276462797076
step: 140, loss: 0.0008164063910953701
step: 150, loss: 0.000393153284676373
step: 160, loss: 0.0001533728209324181
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 170, loss: 0.0002822131209541112
step: 180, loss: 0.003024768317118287
step: 190, loss: 0.0018667856929823756
/usr/local/lib/python3.7/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=0.7142857142857143, f1=0.7313432835820897,
best_f1=0.7506849315068493
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2049it [00:12, 167.92it/s]

```



```
load_f1 = 0.6595744680851063
real_f1 = 0.6351706036745407
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, 189.00it/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: 0.6272090077400208
step: 10, loss: 0.38127267360687256
step: 20, loss: 0.3090824484825134
step: 30, loss: 0.3911212980747223
step: 40, loss: 0.3032062351703644
step: 50, loss: 0.2918131649494171
step: 60, loss: 0.2445298135280609
step: 70, loss: 0.38189730048179626
step: 80, loss: 0.3876264691352844
step: 90, loss: 0.2712785005569458
step: 100, loss: 0.19382286071777344
step: 110, loss: 0.30169758200645447
step: 120, loss: 0.31369197368621826
step: 130, loss: 0.08208607882261276
step: 140, loss: 0.24860744178295135

```

```

step: 150, loss: 0.40352728962898254
step: 160, loss: 0.15627674758434296
step: 170, loss: 0.22119641304016113
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6756032171581768, f1=0.6734693877551021,
best_f1=0.6734693877551021
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14165236055850983
step: 10, loss: 0.22919289767742157
step: 20, loss: 0.27633652091026306
step: 30, loss: 0.21091653406620026
step: 40, loss: 0.0757661685347557
step: 50, loss: 0.1178257092833519
step: 60, loss: 0.14027275145053864
step: 70, loss: 0.07544797658920288
step: 80, loss: 0.07636208832263947
step: 90, loss: 0.10879365354776382
step: 100, loss: 0.1726323366165161
step: 110, loss: 0.051858317106962204
step: 120, loss: 0.09488281607627869
step: 130, loss: 0.12030195444822311
step: 140, loss: 0.22853170335292816
step: 150, loss: 0.2925478219985962
step: 160, loss: 0.16011720895767212
step: 170, loss: 0.09213755279779434
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.756476683937824, f1=0.7680798004987531,
best_f1=0.7680798004987531
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12841688096523285
step: 10, loss: 0.14926093816757202
step: 20, loss: 0.008932141587138176

```

```

step: 30, loss: 0.04640290513634682
step: 40, loss: 0.04843549802899361
step: 50, loss: 0.03518828749656677
step: 60, loss: 0.09173990041017532
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 70, loss: 0.08927835524082184
step: 80, loss: 0.06626749038696289
step: 90, loss: 0.1240176409482956
step: 100, loss: 0.0033750850707292557
step: 110, loss: 0.11365824937820435
step: 120, loss: 0.059336550533771515
step: 130, loss: 0.16666094958782196
step: 140, loss: 0.01857374794781208
step: 150, loss: 0.061540380120277405
step: 160, loss: 0.07326879352331161
step: 170, loss: 0.10716541111469269
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7755102040816327, f1=0.7688679245283018,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.058197248727083206
step: 10, loss: 0.11673532426357269
step: 20, loss: 0.0034445361234247684
step: 30, loss: 0.0760282427072525
step: 40, loss: 0.02377631701529026
step: 50, loss: 0.2253226935863495
step: 60, loss: 0.06033085286617279
step: 70, loss: 0.03072541020810604
step: 80, loss: 0.0828511044383049
step: 90, loss: 0.07655578851699829
step: 100, loss: 0.03879629075527191
step: 110, loss: 0.08079466223716736
step: 120, loss: 0.0030069102067500353
step: 130, loss: 0.07026499509811401
step: 140, loss: 0.03137753903865814
step: 150, loss: 0.21978963911533356
step: 160, loss: 0.11893024295568466
step: 170, loss: 0.04320043697953224
/usr/local/lib/python3.7/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=0.7721822541966428, f1=0.7710280373831775,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028483876958489418
step: 10, loss: 0.13459806144237518
step: 20, loss: 0.019937116652727127
step: 30, loss: 0.09612845629453659
step: 40, loss: 0.015561318024992943
step: 50, loss: 0.13878552615642548
step: 60, loss: 0.021843524649739265
step: 70, loss: 0.1177259236574173
step: 80, loss: 0.012462607584893703
step: 90, loss: 0.0818922370672226
step: 100, loss: 0.03715498745441437
step: 110, loss: 0.04554593190550804
step: 120, loss: 0.006128587760031223
step: 130, loss: 0.0260696429759264
step: 140, loss: 0.04688117280602455
step: 150, loss: 0.013373411260545254
step: 160, loss: 0.03252377733588219
step: 170, loss: 0.004353160038590431
/usr/local/lib/python3.7/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=0.7657430730478589, f1=0.7548076923076923,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002234310144558549
step: 10, loss: 0.006863039452582598
step: 20, loss: 0.0011233389377593994
step: 30, loss: 0.0016708807088434696
step: 40, loss: 0.03227957710623741
step: 50, loss: 0.09357087314128876
step: 60, loss: 0.07103057205677032
step: 70, loss: 0.0867852121591568
step: 80, loss: 0.023564213886857033
step: 90, loss: 0.0015794681385159492
step: 100, loss: 0.0012969280360266566
step: 110, loss: 0.0014416440390050411
step: 120, loss: 0.010114409029483795
step: 130, loss: 0.007782737724483013
step: 140, loss: 0.013663779944181442
step: 150, loss: 0.06868848204612732

```

```

step: 160, loss: 0.025752361863851547
step: 170, loss: 0.02962392382323742
/usr/local/lib/python3.7/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=0.7512953367875648, f1=0.7725118483412323,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014567369362339377
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 10, loss: 0.019895410165190697
step: 20, loss: 0.0014554045628756285
step: 30, loss: 0.017656132578849792
step: 40, loss: 0.04253470525145531
step: 50, loss: 0.12194421142339706
step: 60, loss: 0.005072849337011576
step: 70, loss: 0.009602148085832596
step: 80, loss: 0.01936723291873932
step: 90, loss: 0.00039691090933047235
step: 100, loss: 0.0006932539981789887
step: 110, loss: 0.012117685750126839
step: 120, loss: 0.0007063204539008439
step: 130, loss: 0.12963423132896423
step: 140, loss: 0.004921023268252611
step: 150, loss: 0.002704037819057703
step: 160, loss: 0.0009630953427404165
step: 170, loss: 0.14317521452903748
/usr/local/lib/python3.7/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=0.7464114832535885, f1=0.7499999999999999,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01625184714794159
step: 10, loss: 0.005549589637666941
step: 20, loss: 0.0033612444531172514
step: 30, loss: 0.008439213037490845
step: 40, loss: 0.052495185285806656
step: 50, loss: 0.026440147310495377
step: 60, loss: 0.0072659822180867195
step: 70, loss: 0.0009796220110729337
step: 80, loss: 0.00206630933098495
step: 90, loss: 0.02392684854567051
step: 100, loss: 0.09685068577528

```

```

step: 110, loss: 0.018778035417199135
step: 120, loss: 0.004929743241518736
step: 130, loss: 0.003941280767321587
step: 140, loss: 0.0014663721667602658
step: 150, loss: 0.03704174607992172
step: 160, loss: 0.007713672239333391
step: 170, loss: 0.036262061446905136
/usr/local/lib/python3.7/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=0.7651715039577837, f1=0.7524752475247525,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.012080197222530842
step: 10, loss: 0.004130041692405939
step: 20, loss: 0.001967289485037327
step: 30, loss: 0.013459998182952404
step: 40, loss: 0.0035579949617385864
step: 50, loss: 0.0023912754841148853
step: 60, loss: 0.05203881114721298
step: 70, loss: 0.15002089738845825
step: 80, loss: 0.027365846559405327
step: 90, loss: 0.0586448609828949
step: 100, loss: 0.033907707780599594
step: 110, loss: 0.0011771870777010918
step: 120, loss: 0.002778099151328206
step: 130, loss: 0.07848809659481049
step: 140, loss: 0.00877409614622593
step: 150, loss: 0.0017744106007739902
step: 160, loss: 0.085018590092659
step: 170, loss: 0.01622164063155651
/usr/local/lib/python3.7/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=0.7596153846153846, f1=0.7681818181818182,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10806838423013687
step: 10, loss: 0.0092779491096735
step: 20, loss: 0.023431140929460526
step: 30, loss: 0.0043548764660954475
step: 40, loss: 0.0006638760096393526
step: 50, loss: 0.08074771612882614
step: 60, loss: 0.00036719435593113303

```

```

step: 70, loss: 0.0038476241752505302
step: 80, loss: 0.01433771662414074
step: 90, loss: 0.002760587725788355
step: 100, loss: 0.0008133301744237542
step: 110, loss: 0.004245097283273935
step: 120, loss: 0.01385500468313694
step: 130, loss: 0.004982966464012861
step: 140, loss: 0.05802387744188309
step: 150, loss: 0.007778482977300882
step: 160, loss: 0.0005050771287642419
step: 170, loss: 0.0003119247267022729
/usr/local/lib/python3.7/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=0.7581047381546135, f1=0.7541766109785203,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0197402685880661
step: 10, loss: 0.013437814079225063
step: 20, loss: 0.005156477447599173
step: 30, loss: 0.001062369323335588
step: 40, loss: 0.020047252997756004
step: 50, loss: 0.04601817950606346
step: 60, loss: 0.009608946740627289
step: 70, loss: 0.03704072907567024
step: 80, loss: 0.0005911254556849599
step: 90, loss: 0.00018102095054928213
step: 100, loss: 0.002250737277790904
step: 110, loss: 0.05561346933245659
step: 120, loss: 0.0009517860598862171
step: 130, loss: 0.0003865660692099482
step: 140, loss: 0.0006496772984974086
step: 150, loss: 0.11108864843845367
step: 160, loss: 0.002936396049335599
step: 170, loss: 0.004817601293325424
/usr/local/lib/python3.7/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=0.7556675062972292, f1=0.7523809523809523,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01251536887139082
step: 10, loss: 0.018253641203045845
step: 20, loss: 0.11297564208507538

```



```

step: 30, loss: 0.005339613649994135
step: 40, loss: 0.0004806591023225337
step: 50, loss: 0.00018160806212108582
step: 60, loss: 0.034884318709373474
step: 70, loss: 0.13248953223228455
step: 80, loss: 0.0001534604816697538
step: 90, loss: 0.14134572446346283
step: 100, loss: 0.0037743018474429846
step: 110, loss: 0.0011392944725230336
step: 120, loss: 0.04287010058760643
step: 130, loss: 0.009853308089077473
step: 140, loss: 0.0019428411033004522
step: 150, loss: 0.00028555651078931987
step: 160, loss: 0.00039862070116214454
step: 170, loss: 0.011721797287464142
/usr/local/lib/python3.7/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=0.7595628415300546, f1=0.7671957671957672,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.007945985533297062
step: 10, loss: 0.00035827048122882843
step: 20, loss: 0.006617716513574123
step: 30, loss: 0.00012997680460102856
step: 40, loss: 0.0006516239955089986
step: 50, loss: 0.0004936133045703173
step: 60, loss: 0.0007184789865277708
step: 70, loss: 0.0001804921921575442
step: 80, loss: 0.00016871636034920812
step: 90, loss: 0.00011573176743695512
step: 100, loss: 0.00027511175721883774
step: 110, loss: 0.00012143935600761324
step: 120, loss: 0.018911877647042274
step: 130, loss: 0.0009418231202289462
step: 140, loss: 0.0005375276086851954
step: 150, loss: 0.0024090022780001163
step: 160, loss: 0.00921645388007164
step: 170, loss: 0.0018211203860118985
/usr/local/lib/python3.7/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=0.7500000000000001, f1=0.7680412371134021,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0062810806557536125
step: 10, loss: 8.476847142446786e-05
step: 20, loss: 0.0005069053149782121
step: 30, loss: 0.0008151368238031864
step: 40, loss: 0.00016786908963695168
step: 50, loss: 0.0002784480748232454
step: 60, loss: 0.00012219826749060303
step: 70, loss: 0.00038152013439685106
step: 80, loss: 0.011690761893987656
step: 90, loss: 0.0002992596128024161
step: 100, loss: 0.009593332186341286
step: 110, loss: 0.00034257094375789165
step: 120, loss: 0.0019902389030903578
step: 130, loss: 0.00026013964088633657
step: 140, loss: 0.00035425368696451187
step: 150, loss: 0.00012630656419787556
step: 160, loss: 0.00010131965245818719
step: 170, loss: 0.00016372684331145138
/usr/local/lib/python3.7/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=0.7520435967302452, f1=0.764857881136951,
best_f1=0.7688679245283018
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000180018600076437
step: 10, loss: 0.003104656469076872
step: 20, loss: 0.03168192133307457
step: 30, loss: 0.00014490759349428117
step: 40, loss: 0.0006150020053610206
step: 50, loss: 0.00018530416127759963
step: 60, loss: 0.013515711762011051
step: 70, loss: 0.00157257413957268
step: 80, loss: 0.0031721496488898993
step: 90, loss: 0.006104915868490934
step: 100, loss: 0.001688860240392387
step: 110, loss: 0.00015658498159609735
step: 120, loss: 0.004132910631597042
step: 130, loss: 0.002036832505837083
step: 140, loss: 0.0021338665392249823
step: 150, loss: 0.00021611513511743397
step: 160, loss: 0.00045474700164049864
step: 170, loss: 0.0007549778092652559
/usr/local/lib/python3.7/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=0.7513513513513513, f1=0.7641025641025642,
best_f1=0.7688679245283018
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, 231.66it/s]
```

```
load_f1 = 0.6330935251798562
```

```
real_f1 = 0.6258823529411764
```

```
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, 189.15it/s]
```

3.3 DK DIRTY

3.3.1 DBLP-ACM - Running the matcher

```
[ ]: # Step 1:
# train the wdc model with the save_model flag on
# it will produce two files *_dev.pt and *_test.pt which are the
# best checkpoints on the dev set and the test set.
!CUDA_VISIBLE_DEVICES=0 python train_ditto.py \
--task Dirty/DBLP-ACM \
--batch_size 32 \
--max_len 256 \
--lr 3e-5 \
--n_epochs 15 \
--finetuning \
--lm 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: 0.5942934155464172
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 10, loss: 0.6223584413528442
step: 20, loss: 0.5023674964904785
step: 30, loss: 0.42634424567222595
step: 40, loss: 0.22183197736740112
step: 50, loss: 0.09503340721130371
step: 60, loss: 0.14560693502426147
step: 70, loss: 0.09036318957805634
step: 80, loss: 0.07510967552661896
step: 90, loss: 0.12885206937789917
step: 100, loss: 0.007916983217000961
step: 110, loss: 0.26320669054985046
step: 120, loss: 0.01510709896683693
step: 130, loss: 0.06864400953054428
step: 140, loss: 0.022758392617106438
step: 150, loss: 0.062369246035814285
step: 160, loss: 0.011984096840023994
step: 170, loss: 0.04449733346700668
step: 180, loss: 0.10497231781482697
step: 190, loss: 0.10053414106369019
step: 200, loss: 0.13519041240215302
step: 210, loss: 0.025382397696375847

```

```

step: 220, loss: 0.03234328702092171
step: 230, loss: 0.08869108557701111
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9599999999999999, f1=0.9588431590656283,
best_f1=0.9588431590656283
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.008047888055443764
step: 10, loss: 0.021715428680181503
step: 20, loss: 0.23411186039447784
step: 30, loss: 0.14371611177921295
step: 40, loss: 0.1312599778175354
step: 50, loss: 0.02533729560673237
step: 60, loss: 0.008791483007371426
step: 70, loss: 0.06340094655752182
step: 80, loss: 0.04656525328755379
step: 90, loss: 0.02559984102845192
step: 100, loss: 0.1048821285367012
step: 110, loss: 0.10128708928823471
step: 120, loss: 0.09007735550403595
step: 130, loss: 0.13043129444122314
step: 140, loss: 0.008941681124269962
step: 150, loss: 0.013950178399682045
step: 160, loss: 0.024577265605330467
step: 170, loss: 0.02999546192586422
step: 180, loss: 0.06451800465583801
step: 190, loss: 0.03698594868183136
step: 200, loss: 0.003693557810038328
step: 210, loss: 0.010809729807078838
step: 220, loss: 0.05148041248321533
step: 230, loss: 0.07886763662099838
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9650507328072153, f1=0.9542920847268673,
best_f1=0.9542920847268673
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018564175814390182
step: 10, loss: 0.016285011544823647
step: 20, loss: 0.09982914477586746
step: 30, loss: 0.0024156842846423388
step: 40, loss: 0.011784033849835396
step: 50, loss: 0.04000045731663704
step: 60, loss: 0.0017688007792457938
step: 70, loss: 0.0032028695568442345
step: 80, loss: 0.064728744328022
step: 90, loss: 0.1363748013973236
step: 100, loss: 0.027457214891910553
step: 110, loss: 0.0014480644604191184
step: 120, loss: 0.052246470004320145
step: 130, loss: 0.0005951575003564358
step: 140, loss: 0.007271731738001108
step: 150, loss: 0.11068245023488998
step: 160, loss: 0.23120296001434326
step: 170, loss: 0.014856550842523575
step: 180, loss: 0.025257380679249763
step: 190, loss: 0.008742891252040863
step: 200, loss: 0.07254630327224731
step: 210, loss: 0.018847908824682236
step: 220, loss: 0.04062582924962044
step: 230, loss: 0.008349094539880753
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9741863075196409, f1=0.962962962962963,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0124070905148983
step: 10, loss: 0.0024191592819988728
step: 20, loss: 0.00633185263723135
step: 30, loss: 0.0006634573801420629
step: 40, loss: 0.012174688279628754
step: 50, loss: 0.000765860197134316
step: 60, loss: 0.002392369555309415
step: 70, loss: 0.015490110963582993
step: 80, loss: 0.00045258968020789325
step: 90, loss: 0.060439079999923706

```

```

step: 100, loss: 0.007060652133077383
step: 110, loss: 0.001730823889374733
step: 120, loss: 0.08559790253639221
step: 130, loss: 0.02381112426519394
step: 140, loss: 0.0029316111467778683
step: 150, loss: 0.010115277022123337
step: 160, loss: 0.04323989525437355
step: 170, loss: 0.008360245265066624
step: 180, loss: 0.028247809037566185
step: 190, loss: 0.0007357181166298687
step: 200, loss: 0.002569507574662566
step: 210, loss: 0.009916557930409908
step: 220, loss: 0.0004536872438620776
step: 230, loss: 0.10064787417650223
/usr/local/lib/python3.7/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=0.9720044792833147, f1=0.9684684684684683,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000666029576677829
step: 10, loss: 0.0015127032529562712
step: 20, loss: 0.0013800198212265968
step: 30, loss: 0.00024373845371883363
step: 40, loss: 0.0014881453244015574
step: 50, loss: 0.00046048726653680205
step: 60, loss: 0.06876261532306671
step: 70, loss: 0.0006579623441211879
step: 80, loss: 0.0006181250209920108
step: 90, loss: 0.00020263706392142922
step: 100, loss: 0.00020478165242820978
step: 110, loss: 0.00047492951853200793
step: 120, loss: 0.00011997273395536467
step: 130, loss: 0.0023787971585989
step: 140, loss: 0.0006570280529558659
step: 150, loss: 0.00047113001346588135
step: 160, loss: 0.0008747380925342441
step: 170, loss: 0.036828916519880295
step: 180, loss: 0.0004029624105896801
step: 190, loss: 0.19882087409496307
step: 200, loss: 0.00357147422619164
step: 210, loss: 0.0010289894416928291
step: 220, loss: 0.0010942468652501702
step: 230, loss: 0.0006858743727207184
/usr/local/lib/python3.7/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=0.9709821428571428, f1=0.9608938547486034,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007980498485267162
step: 10, loss: 0.0012260766234248877
step: 20, loss: 0.0020735773723572493
step: 30, loss: 0.006343163549900055
step: 40, loss: 0.00023090261674951762
step: 50, loss: 0.0002955567033495754
step: 60, loss: 0.00017653388204053044
step: 70, loss: 0.0003744979912880808
step: 80, loss: 0.0011075938818976283
step: 90, loss: 0.0006049896473996341
step: 100, loss: 0.01717032492160797
step: 110, loss: 0.000776332279201597
step: 120, loss: 0.002128052758052945
step: 130, loss: 0.0002089262561639771
step: 140, loss: 0.00040168495615944266
step: 150, loss: 0.0005745503003709018
step: 160, loss: 0.0009912626119330525
step: 170, loss: 0.00020998621766921133
step: 180, loss: 0.017933491617441177
step: 190, loss: 0.0008501936681568623
step: 200, loss: 0.00014977519458625466
step: 210, loss: 0.0002674727002158761
step: 220, loss: 0.0001536237687105313
step: 230, loss: 0.038845401257276535
/usr/local/lib/python3.7/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=0.9664429530201343, f1=0.9652855543113102,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.03706645220518112
step: 10, loss: 0.00023800673079676926
step: 20, loss: 0.00013244253932498395
step: 30, loss: 9.92726199910976e-05
step: 40, loss: 0.0060447631403803825
step: 50, loss: 0.00016782748571131378
step: 60, loss: 0.00035985917202197015
step: 70, loss: 0.00021983204351272434
step: 80, loss: 0.0014162735315039754
step: 90, loss: 0.000130749904201366

```

```

step: 100, loss: 0.00013287272304296494
step: 110, loss: 0.0005890750326216221
step: 120, loss: 0.0006773289642296731
step: 130, loss: 0.00010755871335277334
step: 140, loss: 0.0001464174856664613
step: 150, loss: 0.007302244193851948
step: 160, loss: 0.0405389778316021
step: 170, loss: 0.00020586659957189113
step: 180, loss: 0.0001533509057480842
step: 190, loss: 0.00012413671356625855
step: 200, loss: 0.01859789900481701
step: 210, loss: 7.976614142535254e-05
step: 220, loss: 0.0012315824860706925
step: 230, loss: 0.004452583380043507
/usr/local/lib/python3.7/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=0.9666666666666666, f1=0.9633740288568259,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005359696224331856
step: 10, loss: 0.0006390538183040917
step: 20, loss: 0.00019557114865165204
step: 30, loss: 0.00017303938511759043
step: 40, loss: 0.019014369696378708
step: 50, loss: 0.0007108469144441187
step: 60, loss: 0.00016309581405948848
step: 70, loss: 0.0001404736831318587
step: 80, loss: 0.0004273138474673033
step: 90, loss: 0.00011209453805349767
step: 100, loss: 0.00036398094380274415
step: 110, loss: 0.000751428073272109
step: 120, loss: 0.003780290950089693
step: 130, loss: 0.0016711377538740635
step: 140, loss: 0.0001355957647319883
step: 150, loss: 0.0007276323740370572
step: 160, loss: 0.0004772973188664764
step: 170, loss: 0.00014730360999237746
step: 180, loss: 0.0002533066435717046
step: 190, loss: 0.000592519121710211
step: 200, loss: 0.002185328397899866
step: 210, loss: 0.02510712668299675
step: 220, loss: 0.0025308614131063223
step: 230, loss: 0.002061246894299984
/usr/local/lib/python3.7/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=0.9640449438202248, f1=0.9593679458239278,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002202571922680363
step: 10, loss: 0.0008188455831259489
step: 20, loss: 0.000776022847276181
step: 30, loss: 0.0002888769959099591
step: 40, loss: 0.0010925988899543881
step: 50, loss: 0.0006669831927865744
step: 60, loss: 0.00018002190336119384
step: 70, loss: 0.0002340326172998175
step: 80, loss: 0.0002511347702238709
step: 90, loss: 0.0011867471039295197
step: 100, loss: 0.00019173999316990376
step: 110, loss: 9.737860091263428e-05
step: 120, loss: 8.964732842287049e-05
step: 130, loss: 0.002393710194155574
step: 140, loss: 6.0678907175315544e-05
step: 150, loss: 0.0012992055853828788
step: 160, loss: 0.000387549982406199
step: 170, loss: 0.00011323228682158515
step: 180, loss: 0.0033018155954778194
step: 190, loss: 0.00011007816647179425
step: 200, loss: 0.00011687763617374003
step: 210, loss: 0.00031590129947289824
step: 220, loss: 9.516568388789892e-05
step: 230, loss: 6.799647962907329e-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=0.9732142857142857, f1=0.9663677130044843,
best_f1=0.962962962962963
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 9.770996257429942e-05
step: 10, loss: 0.00015383226855192333
step: 20, loss: 5.2385134040378034e-05
step: 30, loss: 0.0005859886878170073
step: 40, loss: 0.0002640803868416697
step: 50, loss: 6.379722617566586e-05
step: 60, loss: 0.00021922189625911415
step: 70, loss: 0.0001878298062365502
step: 80, loss: 8.43964735395275e-05
step: 90, loss: 0.00010763064346974716

```

```

step: 100, loss: 6.932541873538867e-05
step: 110, loss: 0.00023423477250616997
step: 120, loss: 0.00024045137979555875
step: 130, loss: 0.0002820095978677273
step: 140, loss: 0.03456759452819824
step: 150, loss: 0.0008545666350983083
step: 160, loss: 0.0004273036902304739
step: 170, loss: 5.0087302952306345e-05
step: 180, loss: 0.0002840836241375655
step: 190, loss: 0.001154901459813118
step: 200, loss: 7.907598774181679e-05
step: 210, loss: 0.00010789002408273518
step: 220, loss: 5.910839172429405e-05
step: 230, loss: 0.0003718712250702083
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9745293466223698, f1=0.9546961325966851,
best_f1=0.9546961325966851
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002161906595574692
step: 10, loss: 7.936025212984532e-05
step: 20, loss: 0.00015201029600575566
step: 30, loss: 0.018757177516818047
step: 40, loss: 0.0004980018711648881
step: 50, loss: 0.00014101459237281233
step: 60, loss: 0.010260957293212414
step: 70, loss: 0.0002955591189675033
step: 80, loss: 0.00014064053539186716
step: 90, loss: 0.00012906905612908304
step: 100, loss: 8.777618495514616e-05
step: 110, loss: 5.004241756978445e-05
step: 120, loss: 0.0004995919880457222
step: 130, loss: 0.0003139539039693773
step: 140, loss: 6.648148701060563e-05
step: 150, loss: 0.03727380558848381
step: 160, loss: 0.0010196119546890259
step: 170, loss: 0.005667916499078274
step: 180, loss: 0.00013845543435309082
step: 190, loss: 5.258977398625575e-05
step: 200, loss: 0.0004366462235338986
step: 210, loss: 3.446527989581227e-05

```

```

step: 220, loss: 4.228375473758206e-05
step: 230, loss: 6.28934649284929e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 11: dev_f1=0.9787709497206705, f1=0.968609865470852,
best_f1=0.968609865470852
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003645766410045326
step: 10, loss: 5.165278707863763e-05
step: 20, loss: 0.0035341838374733925
step: 30, loss: 0.00024455561651848257
step: 40, loss: 0.0012943159090355039
step: 50, loss: 0.01074350718408823
step: 60, loss: 0.0006506209028884768
step: 70, loss: 0.00018184138752985746
step: 80, loss: 6.386582390405238e-05
step: 90, loss: 3.6219465982867405e-05
step: 100, loss: 5.537614197237417e-05
step: 110, loss: 0.0002802413364406675
step: 120, loss: 4.842174894292839e-05
step: 130, loss: 0.00019441236509010196
step: 140, loss: 8.181680459529161e-05
step: 150, loss: 0.003874568035826087
step: 160, loss: 5.610975495073944e-05
step: 170, loss: 8.104465814540163e-05
step: 180, loss: 8.620653534308076e-05
step: 190, loss: 0.0008349997806362808
step: 200, loss: 4.367953079054132e-05
step: 210, loss: 4.765084304381162e-05
step: 220, loss: 4.029836054542102e-05
step: 230, loss: 0.00044734939001500607
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9810479375696767, f1=0.967525195968645,
best_f1=0.967525195968645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.

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

```
step: 0, loss: 0.0013293352676555514
step: 10, loss: 0.003332240507006645
step: 20, loss: 0.0003444541071075946
step: 30, loss: 0.0007834902498871088
step: 40, loss: 0.00015816903032828122
step: 50, loss: 7.98096225480549e-05
step: 60, loss: 0.0002878580125980079
step: 70, loss: 6.91498935339041e-05
step: 80, loss: 9.552847768645734e-05
step: 90, loss: 0.005357328802347183
step: 100, loss: 3.763879794860259e-05
step: 110, loss: 9.629586566006765e-05
step: 120, loss: 0.022586364299058914
step: 130, loss: 9.087083162739873e-05
step: 140, loss: 6.711656169500202e-05
step: 150, loss: 5.361109651857987e-05
step: 160, loss: 5.247794615570456e-05
step: 170, loss: 0.00022686709417030215
step: 180, loss: 5.108597542857751e-05
step: 190, loss: 6.211180152604356e-05
step: 200, loss: 0.004735888447612524
step: 210, loss: 3.294894122518599e-05
step: 220, loss: 0.00011373167217243463
step: 230, loss: 7.882415229687467e-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=0.9776785714285714, f1=0.9652076318742986,
best_f1=0.967525195968645
```

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

An input tensor was not cuda.

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

```
step: 0, loss: 6.255327753024176e-05
step: 10, loss: 0.0022861389443278313
step: 20, loss: 8.849816367728636e-05
step: 30, loss: 5.2098563173785806e-05
step: 40, loss: 0.00016896906890906394
step: 50, loss: 8.075936057139188e-05
step: 60, loss: 0.0005939262337051332
step: 70, loss: 4.444065780262463e-05
step: 80, loss: 6.375920202117413e-05
step: 90, loss: 0.00010305096657248214
step: 100, loss: 4.9539219617145136e-05
step: 110, loss: 0.001420677057467401
step: 120, loss: 3.454606849118136e-05
step: 130, loss: 6.735907663824037e-05
```

```

step: 140, loss: 0.00043236996862106025
step: 150, loss: 0.00010836866567842662
step: 160, loss: 0.006388865411281586
step: 170, loss: 2.491800842108205e-05
step: 180, loss: 7.182966510299593e-05
step: 190, loss: 7.004093640716746e-05
step: 200, loss: 3.900913361576386e-05
step: 210, loss: 0.00017443728575017303
step: 220, loss: 0.00020356416644062847
step: 230, loss: 0.00015541860193479806
/usr/local/lib/python3.7/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=0.9743589743589743, f1=0.9640449438202248,
best_f1=0.967525195968645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 3.142165587632917e-05
step: 10, loss: 3.041549825866241e-05
step: 20, loss: 0.0001162485932582058
step: 30, loss: 8.959570550359786e-05
step: 40, loss: 7.263280713232234e-05
step: 50, loss: 0.019346781075000763
step: 60, loss: 9.126860095420852e-05
step: 70, loss: 4.580732638714835e-05
step: 80, loss: 8.016022184165195e-05
step: 90, loss: 5.485156725626439e-05
step: 100, loss: 4.80494272778742e-05
step: 110, loss: 5.3298153943615034e-05
step: 120, loss: 0.00010598717926768586
step: 130, loss: 5.5599633924430236e-05
step: 140, loss: 4.6878692955942824e-05
step: 150, loss: 0.0001240180863533169
step: 160, loss: 3.2941468816716224e-05
step: 170, loss: 4.439753683982417e-05
step: 180, loss: 5.978147964924574e-05
step: 190, loss: 7.696681859670207e-05
step: 200, loss: 7.078200724208727e-05
step: 210, loss: 0.00012305534619372338
step: 220, loss: 4.7646884922869503e-05
step: 230, loss: 4.8842932301340625e-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=0.9754464285714286, f1=0.9640449438202248,
best_f1=0.967525195968645
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, 187.53it/s]

load_f1 = 0.9720044792833147

real_f1 = 0.9665178571428571

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, 230.19it/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 02: FP16 training with FP32 batchnorm and FP32 master weights.

Defaults for this optimization level are:

```

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

```

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

After processing overrides, optimization options are:

```

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

```

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

```

Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/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: 0.6143800020217896
step: 10, loss: 0.6107079982757568
step: 20, loss: 0.5495185852050781
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 30, loss: 0.3344217538833618
step: 40, loss: 0.2724335193634033
step: 50, loss: 0.2885284125804901
step: 60, loss: 0.06977561861276627
step: 70, loss: 0.22178246080875397
step: 80, loss: 0.07152446359395981
step: 90, loss: 0.5112090110778809
step: 100, loss: 0.09097181260585785
step: 110, loss: 0.10768565535545349
step: 120, loss: 0.25375062227249146
step: 130, loss: 0.1068514809012413
step: 140, loss: 0.07895254343748093
step: 150, loss: 0.0529199093580246
step: 160, loss: 0.04370030388236046
step: 170, loss: 0.27685093879699707
step: 180, loss: 0.24286963045597076
step: 190, loss: 0.01395308505743742
step: 200, loss: 0.18746314942836761
step: 210, loss: 0.11097604036331177
step: 220, loss: 0.26108336448669434
step: 230, loss: 0.1655369997024536
step: 240, loss: 0.11037975549697876
step: 250, loss: 0.06033053621649742
step: 260, loss: 0.18583513796329498
step: 270, loss: 0.04364939406514168
step: 280, loss: 0.09031523019075394
step: 290, loss: 0.24646110832691193
step: 300, loss: 0.1088217943906784
step: 310, loss: 0.19477808475494385
step: 320, loss: 0.15098914504051208

```

```

step: 330, loss: 0.044062454253435135
step: 340, loss: 0.03838520869612694
step: 350, loss: 0.14836996793746948
step: 360, loss: 0.05935284495353699
step: 370, loss: 0.09904457628726959
step: 380, loss: 0.016084151342511177
step: 390, loss: 0.23482616245746613
step: 400, loss: 0.2522397041320801
step: 410, loss: 0.04074055328965187
step: 420, loss: 0.12202010303735733
step: 430, loss: 0.1557614803314209
step: 440, loss: 0.041242651641368866
step: 450, loss: 0.014418542385101318
step: 460, loss: 0.026691723614931107
step: 470, loss: 0.12599655985832214
step: 480, loss: 0.07886381447315216
step: 490, loss: 0.07930061221122742
step: 500, loss: 0.04353094473481178
step: 510, loss: 0.0962212011218071
step: 520, loss: 0.05105546861886978
step: 530, loss: 0.007637709379196167
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9183387774148392, f1=0.9160092807424594,
best_f1=0.9160092807424594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14045186340808868
step: 10, loss: 0.13468210399150848
step: 20, loss: 0.06261622160673141
step: 30, loss: 0.06010819599032402
step: 40, loss: 0.052910350263118744
step: 50, loss: 0.23315049707889557
step: 60, loss: 0.0707065761089325
step: 70, loss: 0.055115751922130585
step: 80, loss: 0.05437713861465454
step: 90, loss: 0.026672013103961945
step: 100, loss: 0.09207796305418015
step: 110, loss: 0.047013040632009506
step: 120, loss: 0.2154548615217209
step: 130, loss: 0.06390736252069473
step: 140, loss: 0.06325040012598038

```

```

step: 150, loss: 0.0761009082198143
step: 160, loss: 0.10553694516420364
step: 170, loss: 0.11634659022092819
step: 180, loss: 0.010815311223268509
step: 190, loss: 0.07330746203660965
step: 200, loss: 0.08418290317058563
step: 210, loss: 0.06821706146001816
step: 220, loss: 0.08542603254318237
step: 230, loss: 0.009576880373060703
step: 240, loss: 0.1058693379163742
step: 250, loss: 0.04749695584177971
step: 260, loss: 0.014094185084104538
step: 270, loss: 0.19529592990875244
step: 280, loss: 0.15432611107826233
step: 290, loss: 0.03387221321463585
step: 300, loss: 0.1512218564748764
step: 310, loss: 0.01096223946660757
step: 320, loss: 0.23378074169158936
step: 330, loss: 0.03885616362094879
step: 340, loss: 0.03851478919386864
step: 350, loss: 0.0049616196192801
step: 360, loss: 0.05618948116898537
step: 370, loss: 0.23980532586574554
step: 380, loss: 0.03443773090839386
step: 390, loss: 0.04541506618261337
step: 400, loss: 0.00583788612857461
step: 410, loss: 0.04063829034566879
step: 420, loss: 0.013960538432002068
step: 430, loss: 0.04758969321846962
step: 440, loss: 0.11626311391592026
step: 450, loss: 0.13872846961021423
step: 460, loss: 0.024087101221084595
step: 470, loss: 0.015539290383458138
step: 480, loss: 0.1765856146812439
step: 490, loss: 0.027657093480229378
step: 500, loss: 0.2511291801929474
step: 510, loss: 0.041788164526224136
step: 520, loss: 0.08713793009519577
step: 530, loss: 0.041429970413446426
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9235048678720444, f1=0.9227230910763569,
best_f1=0.9227230910763569

```

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

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

```
step: 0, loss: 0.0794174000620842  
step: 10, loss: 0.0912197157740593  
step: 20, loss: 0.057686809450387955  
step: 30, loss: 0.04591534659266472  
step: 40, loss: 0.029309462755918503  
step: 50, loss: 0.11361280828714371  
step: 60, loss: 0.053782977163791656  
step: 70, loss: 0.009070180356502533  
step: 80, loss: 0.0017217075219377875  
step: 90, loss: 0.006902695633471012  
step: 100, loss: 0.008071749471127987  
step: 110, loss: 0.005535590462386608  
step: 120, loss: 0.003941611386835575  
step: 130, loss: 0.0025418216828256845  
step: 140, loss: 0.030540600419044495  
step: 150, loss: 0.06942742317914963  
step: 160, loss: 0.005730052012950182  
step: 170, loss: 0.07616551965475082  
step: 180, loss: 0.023010922595858574  
step: 190, loss: 0.013406168669462204  
step: 200, loss: 0.011774379760026932  
step: 210, loss: 0.048392027616500854  
step: 220, loss: 0.31706616282463074  
step: 230, loss: 0.09928911924362183  
step: 240, loss: 0.04009568691253662  
step: 250, loss: 0.031692542135715485  
step: 260, loss: 0.04933860898017883  
step: 270, loss: 0.02719922922551632  
step: 280, loss: 0.20206716656684875  
step: 290, loss: 0.02364901639521122  
step: 300, loss: 0.02975475788116455  
step: 310, loss: 0.0028214443009346724  
step: 320, loss: 0.028989462181925774  
step: 330, loss: 0.011692251078784466  
step: 340, loss: 0.038967255502939224  
step: 350, loss: 0.08028235286474228  
step: 360, loss: 0.041403740644454956  
step: 370, loss: 0.04511276260018349  
step: 380, loss: 0.07194174826145172  
step: 390, loss: 0.02014973573386669  
step: 400, loss: 0.06849836558103561  
step: 410, loss: 0.007739088963717222  
step: 420, loss: 0.21488840878009796  
step: 430, loss: 0.10481742769479752  
step: 440, loss: 0.027768755331635475
```

```

step: 450, loss: 0.058370813727378845
step: 460, loss: 0.13656499981880188
step: 470, loss: 0.1537276953458786
step: 480, loss: 0.03513564169406891
step: 490, loss: 0.002395092276856303
step: 500, loss: 0.13284389674663544
step: 510, loss: 0.05093395337462425
step: 520, loss: 0.1444637030363083
step: 530, loss: 0.15148626267910004
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9243697478991597, f1=0.9187935034802784,
best_f1=0.9187935034802784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.08470093458890915
step: 10, loss: 0.0034293034113943577
step: 20, loss: 0.0036531619261950254
step: 30, loss: 0.005986770614981651
step: 40, loss: 0.035343170166015625
step: 50, loss: 0.0024830952752381563
step: 60, loss: 0.004348542541265488
step: 70, loss: 0.00917125679552555
step: 80, loss: 0.0703708678483963
step: 90, loss: 0.21332202851772308
step: 100, loss: 0.005757179576903582
step: 110, loss: 0.012264780700206757
step: 120, loss: 0.002559575717896223
step: 130, loss: 0.008404947817325592
step: 140, loss: 0.007065438665449619
step: 150, loss: 0.053603094071149826
step: 160, loss: 0.05154614523053169
step: 170, loss: 0.039344679564237595
step: 180, loss: 0.004451955668628216
step: 190, loss: 0.014134763740003109
step: 200, loss: 0.011736392043530941
step: 210, loss: 0.06970918923616409
step: 220, loss: 0.0070346323773264885
step: 230, loss: 0.05760076642036438
step: 240, loss: 0.009147468954324722
step: 250, loss: 0.015964161604642868
step: 260, loss: 0.009704791940748692

```

```

step: 270, loss: 0.009735665284097195
step: 280, loss: 0.04866034910082817
step: 290, loss: 0.08362667262554169
step: 300, loss: 0.0007186868460848927
step: 310, loss: 0.0026541040278971195
step: 320, loss: 0.0029841274954378605
step: 330, loss: 0.002260664477944374
step: 340, loss: 0.04167194664478302
step: 350, loss: 0.02869049645960331
step: 360, loss: 0.03824890777468681
step: 370, loss: 0.007616096641868353
step: 380, loss: 0.04708746075630188
step: 390, loss: 0.0008733168360777199
step: 400, loss: 0.005320468917489052
step: 410, loss: 0.00224745599552989
step: 420, loss: 0.037555016577243805
step: 430, loss: 0.1941710114479065
step: 440, loss: 0.01157333143055439
step: 450, loss: 0.0014886100543662906
step: 460, loss: 0.00202228338457644
step: 470, loss: 0.006266443058848381
step: 480, loss: 0.022444292902946472
step: 490, loss: 0.016017859801650047
step: 500, loss: 0.0031992383301258087
step: 510, loss: 0.03667314723134041
step: 520, loss: 0.03291499987244606
step: 530, loss: 0.04815642535686493
/usr/local/lib/python3.7/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=0.9238532110091744, f1=0.9187214611872146,
best_f1=0.9187935034802784
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04118892922997475
step: 10, loss: 0.03832879662513733
step: 20, loss: 0.004688102286309004
step: 30, loss: 0.003060600720345974
step: 40, loss: 0.019676445052027702
step: 50, loss: 0.016627419739961624
step: 60, loss: 0.010351854376494884
step: 70, loss: 0.002457307418808341
step: 80, loss: 0.00044279469875618815
step: 90, loss: 0.07375013828277588
step: 100, loss: 0.0009567334200255573
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 110, loss: 0.0014463256811723113

```

```
step: 120, loss: 0.02276797406375408
step: 130, loss: 0.0009370458428747952
step: 140, loss: 0.023840108886361122
step: 150, loss: 0.04772519692778587
step: 160, loss: 0.009607616811990738
step: 170, loss: 0.009491724893450737
step: 180, loss: 0.0022538460325449705
step: 190, loss: 0.0032441888470202684
step: 200, loss: 0.03602668642997742
step: 210, loss: 0.0026468115393072367
step: 220, loss: 0.045459579676389694
step: 230, loss: 0.008501755073666573
step: 240, loss: 0.0017420969670638442
step: 250, loss: 0.0040196748450398445
step: 260, loss: 0.08224219083786011
step: 270, loss: 0.0004794928536284715
step: 280, loss: 0.002031089272350073
step: 290, loss: 0.21573208272457123
step: 300, loss: 0.017711395397782326
step: 310, loss: 0.0004294169775675982
step: 320, loss: 0.12062632292509079
step: 330, loss: 0.03956364840269089
step: 340, loss: 0.00544760562479496
step: 350, loss: 0.00575025612488389
step: 360, loss: 0.059293631464242935
step: 370, loss: 0.03280239924788475
step: 380, loss: 0.01659110002219677
step: 390, loss: 0.000406081962864846
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 8192.0
step: 400, loss: 0.01091269962489605
step: 410, loss: 0.012567062862217426
step: 420, loss: 0.00287185562774539
step: 430, loss: 0.05890943482518196
step: 440, loss: 0.05951344966888428
step: 450, loss: 0.011608695611357689
step: 460, loss: 0.001254475093446672
step: 470, loss: 0.05224082991480827
step: 480, loss: 0.0035067580174654722
step: 490, loss: 0.0029539563693106174
step: 500, loss: 0.017411770299077034
step: 510, loss: 0.04000784829258919
step: 520, loss: 0.01823197491466999
step: 530, loss: 0.005805218126624823
/usr/local/lib/python3.7/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=0.8967867575462514, f1=0.8952752070141257,
best_f1=0.9187935034802784
```


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

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

step: 0, loss: 0.001303438562899828
step: 10, loss: 0.0026271785609424114
step: 20, loss: 0.09054131805896759
step: 30, loss: 0.000871316238772124
step: 40, loss: 0.0006168611580505967
step: 50, loss: 0.043248821049928665
step: 60, loss: 0.00038178646354936063
step: 70, loss: 0.010745739564299583
step: 80, loss: 0.0022684696596115828
step: 90, loss: 0.003440654370933771
step: 100, loss: 0.08381766825914383
step: 110, loss: 0.04273701831698418
step: 120, loss: 0.0001499260397395119
step: 130, loss: 0.0019656959921121597
step: 140, loss: 0.0025565146934241056
step: 150, loss: 0.011122038587927818
step: 160, loss: 0.010983188636600971
step: 170, loss: 0.028684202581644058
step: 180, loss: 0.0007835359429009259
step: 190, loss: 0.03759799525141716
step: 200, loss: 0.0003010336949955672
step: 210, loss: 0.014738277532160282
step: 220, loss: 0.0009390326449647546
step: 230, loss: 0.0023324869107455015
step: 240, loss: 0.024812476709485054
step: 250, loss: 0.0028423392213881016
step: 260, loss: 0.009309235028922558
step: 270, loss: 0.1722600758075714
step: 280, loss: 0.0013405239442363381
step: 290, loss: 0.0011463250266388059
step: 300, loss: 0.008097988553345203
step: 310, loss: 0.026912400498986244
step: 320, loss: 0.004690709989517927
step: 330, loss: 0.0015001374995335937
step: 340, loss: 0.13336491584777832
step: 350, loss: 0.004049505572766066
step: 360, loss: 0.0019188944716006517
step: 370, loss: 0.0031061198096722364
step: 380, loss: 0.002074842806905508
step: 390, loss: 0.04850439354777336
step: 400, loss: 0.0002571083896327764
step: 410, loss: 0.014434152282774448
step: 420, loss: 0.002484304364770651
step: 430, loss: 0.00634766323491931
step: 440, loss: 0.00010047991963801906

```

step: 450, loss: 0.0002439011150272563
step: 460, loss: 0.00042983138700947165
step: 470, loss: 0.01407709065824747
step: 480, loss: 0.027520904317498207
step: 490, loss: 0.001890208455733955
step: 500, loss: 0.0034063286148011684
step: 510, loss: 0.03623661771416664
step: 520, loss: 0.0035270731896162033
step: 530, loss: 0.06521754711866379
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9251510925151093, f1=0.9161290322580645,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.006660466082394123
step: 10, loss: 0.00392266595736146
step: 20, loss: 0.0016389790689572692
step: 30, loss: 0.010404354892671108
step: 40, loss: 0.001154205296188593
step: 50, loss: 0.0007050023414194584
step: 60, loss: 0.0014506984734907746
step: 70, loss: 0.045148082077503204
step: 80, loss: 0.02110670506954193
step: 90, loss: 0.02400854043662548
step: 100, loss: 0.00016791837697383016
step: 110, loss: 0.0019041572231799364
step: 120, loss: 0.0006936682038940489
step: 130, loss: 0.0064859576523303986
step: 140, loss: 0.0005775102181360126
step: 150, loss: 0.0006007289630360901
step: 160, loss: 0.0006833840743638575
step: 170, loss: 0.00032431542058475316
step: 180, loss: 0.0002907621383201331
step: 190, loss: 0.0002558463893365115
step: 200, loss: 0.00011421266390243545
step: 210, loss: 0.03459353744983673
step: 220, loss: 0.0003069808299187571
step: 230, loss: 0.0035799795296043158
step: 240, loss: 0.0005277845775708556
step: 250, loss: 0.024713050574064255
step: 260, loss: 0.0002804416581057012

```

```

step: 270, loss: 0.004973102826625109
step: 280, loss: 0.000253889273153618
step: 290, loss: 0.007655272260308266
step: 300, loss: 0.015705937519669533
step: 310, loss: 0.00040564429946243763
step: 320, loss: 0.00015116625581867993
step: 330, loss: 0.0009286087588407099
step: 340, loss: 0.038451388478279114
step: 350, loss: 0.0008159350836649537
step: 360, loss: 0.0028280112892389297
step: 370, loss: 0.0006837784894742072
step: 380, loss: 0.0016764963511377573
step: 390, loss: 0.0003946211945731193
step: 400, loss: 0.004907434806227684
step: 410, loss: 0.011858836747705936
step: 420, loss: 0.0017632985254749656
step: 430, loss: 0.005514903925359249
step: 440, loss: 0.0010220559779554605
step: 450, loss: 0.0006940469611436129
step: 460, loss: 0.00012972527474630624
step: 470, loss: 0.013773364014923573
step: 480, loss: 0.0038199143018573523
step: 490, loss: 0.001555086812004447
step: 500, loss: 0.000590239476878196
step: 510, loss: 0.002915311139076948
step: 520, loss: 0.024404587224125862
step: 530, loss: 0.001074276166036725
/usr/local/lib/python3.7/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=0.9180176007410837, f1=0.9160092807424594,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0013401573523879051
step: 10, loss: 0.0051147048361599445
step: 20, loss: 0.00014262455806601793
step: 30, loss: 0.0005918109090998769
step: 40, loss: 0.0002980410063173622
step: 50, loss: 0.03454600274562836
step: 60, loss: 0.05071518197655678
step: 70, loss: 0.00012477763812057674
step: 80, loss: 0.0002470381441526115
step: 90, loss: 0.00032830442069098353
step: 100, loss: 0.00017701010801829398
step: 110, loss: 0.0004814887070097029
step: 120, loss: 0.000602392596192658

```

```

step: 130, loss: 0.0002552592195570469
step: 140, loss: 0.0005699215107597411
step: 150, loss: 0.0004752445383928716
step: 160, loss: 0.002305036410689354
step: 170, loss: 0.004342278931289911
step: 180, loss: 8.050599717535079e-05
step: 190, loss: 0.004250559490174055
step: 200, loss: 0.0004446472448762506
step: 210, loss: 0.05942107364535332
step: 220, loss: 0.0018906358163803816
step: 230, loss: 0.00036484733573161066
step: 240, loss: 0.0003395344829186797
step: 250, loss: 0.000321995175909251
step: 260, loss: 0.000379120116122067
step: 270, loss: 0.011813885532319546
step: 280, loss: 0.026912132278084755
step: 290, loss: 0.0002750580315478146
step: 300, loss: 0.0003593215951696038
step: 310, loss: 0.005581999663263559
step: 320, loss: 0.0015324440319091082
step: 330, loss: 0.0036529293283820152
step: 340, loss: 0.011417010799050331
step: 350, loss: 0.001868985127657652
step: 360, loss: 0.00462953420355916
step: 370, loss: 0.00011483146954560652
step: 380, loss: 0.0008160565630532801
step: 390, loss: 0.0012284774566069245
step: 400, loss: 0.00019563913519959897
step: 410, loss: 7.622369594173506e-05
step: 420, loss: 0.0003940134774893522
step: 430, loss: 0.009035359136760235
step: 440, loss: 0.0011971023632213473
step: 450, loss: 8.761665958445519e-05
step: 460, loss: 0.0004204483120702207
step: 470, loss: 0.00014276454749051481
step: 480, loss: 7.624840509379283e-05
step: 490, loss: 0.10860516130924225
step: 500, loss: 0.0007669830229133368
step: 510, loss: 0.0001962861279025674
step: 520, loss: 0.0025009089149534702
step: 530, loss: 0.0027319223154336214
/usr/local/lib/python3.7/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=0.913744075829384, f1=0.9133112269066792,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023980385158210993
step: 10, loss: 0.00023668617359362543
step: 20, loss: 0.0012830188497900963
step: 30, loss: 0.0004890567506663501
step: 40, loss: 0.0001194758660858497
step: 50, loss: 0.00010590861347736791
step: 60, loss: 0.0001264856255147606
step: 70, loss: 0.11610980331897736
step: 80, loss: 0.003671110374853015
step: 90, loss: 0.00024385721189901233
step: 100, loss: 0.0002390186709817499
step: 110, loss: 0.0004004244692623615
step: 120, loss: 0.014296798966825008
step: 130, loss: 0.001184467226266861
step: 140, loss: 0.00036267019459046423
step: 150, loss: 0.0013098119525238872
step: 160, loss: 0.00021152480621822178
step: 170, loss: 0.00188922428060323
step: 180, loss: 6.710253364872187e-05
step: 190, loss: 0.00026610997156240046
step: 200, loss: 0.000641010410618037
step: 210, loss: 0.00015001538849901408
step: 220, loss: 0.0005229744710959494
step: 230, loss: 0.0005955936503596604
step: 240, loss: 0.00013739043788518757
step: 250, loss: 8.948751928983256e-05
step: 260, loss: 0.007757789921015501
step: 270, loss: 7.387929508695379e-05
step: 280, loss: 0.04322977364063263
step: 290, loss: 0.0015279370127245784
step: 300, loss: 0.00010845921497093514
step: 310, loss: 0.00025151800946332514
step: 320, loss: 0.00040445144986733794
step: 330, loss: 0.0001508282293798402
step: 340, loss: 0.00010723431478254497
step: 350, loss: 0.00010587657743599266
step: 360, loss: 0.0006168283871375024
step: 370, loss: 0.0003484991320874542
step: 380, loss: 0.00010364697664044797
step: 390, loss: 5.0102487875847146e-05
step: 400, loss: 0.0168928150087595
step: 410, loss: 0.00013247101742308587
step: 420, loss: 4.749779327539727e-05
step: 430, loss: 0.00018816784722730517
step: 440, loss: 0.002808762714266777
step: 450, loss: 0.00011582692968659103
step: 460, loss: 0.0001217479511979036
```

```

step: 470, loss: 0.00016779064026195556
step: 480, loss: 0.0008025650167837739
step: 490, loss: 0.0019291017670184374
step: 500, loss: 0.00041723731555975974
step: 510, loss: 0.0003696575586218387
step: 520, loss: 0.00010395538265584037
step: 530, loss: 0.000744465971365571
/usr/local/lib/python3.7/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=0.9192886456908345, f1=0.915085817524842,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.39518111711368e-05
step: 10, loss: 0.0002313431177753955
step: 20, loss: 0.000128546409541741
step: 30, loss: 0.0002199810405727476
step: 40, loss: 0.0013255588710308075
step: 50, loss: 0.0019189317245036364
step: 60, loss: 0.00018118511070497334
step: 70, loss: 0.000464934651972726
step: 80, loss: 7.151676254579797e-05
step: 90, loss: 0.0006183706573210657
step: 100, loss: 5.399406290962361e-05
step: 110, loss: 5.39096217835322e-05
step: 120, loss: 8.738986070966348e-05
step: 130, loss: 6.409535126294941e-05
step: 140, loss: 0.0016151579329743981
step: 150, loss: 8.026765135582536e-05
step: 160, loss: 0.00035866396501660347
step: 170, loss: 0.01865571178495884
step: 180, loss: 0.004299794789403677
step: 190, loss: 0.0007758447318337858
step: 200, loss: 0.00033166076173074543
step: 210, loss: 0.00847176555544138
step: 220, loss: 0.08945848047733307
step: 230, loss: 0.0011531367199495435
step: 240, loss: 0.0004439821350388229
step: 250, loss: 0.01578388549387455
step: 260, loss: 0.0036296071484684944
step: 270, loss: 0.00012406350288074464
step: 280, loss: 0.0001766445057000965
step: 290, loss: 0.00019983485981356353
step: 300, loss: 8.82840613485314e-05
step: 310, loss: 0.00011363825615262613
step: 320, loss: 0.0008197837742045522

```

```

step: 330, loss: 0.0005792459705844522
step: 340, loss: 0.0001500110374763608
step: 350, loss: 0.0005274183349683881
step: 360, loss: 0.00016257820243481547
step: 370, loss: 0.0012607049429789186
step: 380, loss: 0.0003271974273957312
step: 390, loss: 0.009399236179888248
step: 400, loss: 0.001438653445802629
step: 410, loss: 0.0048252977430820465
step: 420, loss: 0.00287920655682683
step: 430, loss: 6.388821202563122e-05
step: 440, loss: 0.0008165038307197392
step: 450, loss: 0.0003140467160847038
step: 460, loss: 0.0001670285128057003
step: 470, loss: 0.0001284251338802278
step: 480, loss: 4.876195453107357e-05
step: 490, loss: 0.00012084388436051086
step: 500, loss: 0.007599333301186562
step: 510, loss: 0.00017207404016517103
step: 520, loss: 0.000479098001960665
step: 530, loss: 9.863329614745453e-05
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 10: dev_f1=0.9240093240093239, f1=0.9089227924179382,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001887415419332683
step: 10, loss: 0.0002217032015323639
step: 20, loss: 9.840155689744279e-05
step: 30, loss: 0.0003094712446909398
step: 40, loss: 7.682076829951257e-05
step: 50, loss: 0.0001220640042447485
step: 60, loss: 0.0011592821683734655
step: 70, loss: 6.526173820020631e-05
step: 80, loss: 3.8242233131313697e-05
step: 90, loss: 0.0019188262522220612
step: 100, loss: 4.536036794888787e-05
step: 110, loss: 0.0001032659929478541
step: 120, loss: 4.819959576707333e-05
step: 130, loss: 0.0018791857874020934
step: 140, loss: 4.29199353675358e-05
step: 150, loss: 3.689395452965982e-05
step: 160, loss: 0.0002740125637501478
step: 170, loss: 3.673013998195529e-05
step: 180, loss: 0.00037931549013592303

```

```

step: 190, loss: 0.00010959277278743684
step: 200, loss: 0.21925880014896393
step: 210, loss: 0.0015664221718907356
step: 220, loss: 0.0009102922631427646
step: 230, loss: 0.00014119800471235067
step: 240, loss: 0.0004701896687038243
step: 250, loss: 0.0001868308027042076
step: 260, loss: 5.048307139077224e-05
step: 270, loss: 0.012140827253460884
step: 280, loss: 0.003655750071629882
step: 290, loss: 5.198937651584856e-05
step: 300, loss: 4.607940718415193e-05
step: 310, loss: 0.000231081634410657
step: 320, loss: 0.0001273336965823546
step: 330, loss: 0.00015929387882351875
step: 340, loss: 0.002118233125656843
step: 350, loss: 6.834908708697185e-05
step: 360, loss: 0.00011000027734553441
step: 370, loss: 0.0002841016976162791
step: 380, loss: 4.506995901465416e-05
step: 390, loss: 3.8186168239917606e-05
step: 400, loss: 6.474372639786452e-05
step: 410, loss: 7.640023250132799e-05
step: 420, loss: 8.71533848112449e-05
step: 430, loss: 0.00022111473663244396
step: 440, loss: 8.886534487828612e-05
step: 450, loss: 4.3457013816805556e-05
step: 460, loss: 0.0008243880001828074
step: 470, loss: 4.748003993881866e-05
step: 480, loss: 9.653170127421618e-05
step: 490, loss: 0.00025370370713062584
step: 500, loss: 9.129189129453152e-05
step: 510, loss: 0.00038582709385082126
step: 520, loss: 6.521683098981157e-05
step: 530, loss: 0.0002895719662774354
/usr/local/lib/python3.7/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=0.9200743494423792, f1=0.9122969837587007,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0027506486512720585
step: 10, loss: 4.5660726755158976e-05
step: 20, loss: 8.77981074154377e-05
step: 30, loss: 0.005845976062119007
step: 40, loss: 0.00020839097851421684

```


step: 50, loss: 0.010091094300150871
step: 60, loss: 0.0003538182645570487
step: 70, loss: 0.00010746590123744681
step: 80, loss: 5.568841879721731e-05
step: 90, loss: 3.894961628247984e-05
step: 100, loss: 8.93782707862556e-05
step: 110, loss: 0.00016224279534071684
step: 120, loss: 5.6278113333974034e-05
step: 130, loss: 7.37637747079134e-05
step: 140, loss: 4.9596012104302645e-05
step: 150, loss: 3.584473597584292e-05
step: 160, loss: 2.9067054128972813e-05
step: 170, loss: 9.475155093241483e-05
step: 180, loss: 7.131074380595237e-05
step: 190, loss: 0.00046693364856764674
step: 200, loss: 3.6301837099017575e-05
step: 210, loss: 5.889755266252905e-05
step: 220, loss: 0.0001548451546113938
step: 230, loss: 2.5137644115602598e-05
step: 240, loss: 8.298030297737569e-05
step: 250, loss: 0.00021358547382988036
step: 260, loss: 0.029007796198129654
step: 270, loss: 7.082497177179903e-05
step: 280, loss: 5.5576441809535027e-05
step: 290, loss: 5.5403470469173044e-05
step: 300, loss: 0.00103416305501014
step: 310, loss: 0.00010142081009689718
step: 320, loss: 0.027041980996727943
step: 330, loss: 3.0147426514304243e-05
step: 340, loss: 3.139574982924387e-05
step: 350, loss: 0.0001457954931538552
step: 360, loss: 9.114255954045802e-05
step: 370, loss: 0.00031529658008366823
step: 380, loss: 0.00037083119968883693
step: 390, loss: 0.0002743860241025686
step: 400, loss: 6.897807179484516e-05
step: 410, loss: 9.553177369525656e-05
step: 420, loss: 0.0016009691171348095
step: 430, loss: 0.0003329040773678571
step: 440, loss: 6.855140964034945e-05
step: 450, loss: 0.00017616462719161063
step: 460, loss: 3.1410429073730484e-05
step: 470, loss: 3.893806569976732e-05
step: 480, loss: 0.00016966943803709
step: 490, loss: 0.00011057006486225873
step: 500, loss: 0.0005201826570555568
step: 510, loss: 0.00475282734259963
step: 520, loss: 4.74271691928152e-05

```

step: 530, loss: 3.772848504013382e-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=0.9240622140896615, f1=0.9203459262630859,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0003532032133080065
step: 10, loss: 4.681012796936557e-05
step: 20, loss: 0.0003000826691277325
step: 30, loss: 0.002957836724817753
step: 40, loss: 6.570079131051898e-05
step: 50, loss: 0.0002389862056588754
step: 60, loss: 6.443647725973278e-05
step: 70, loss: 3.803774234256707e-05
step: 80, loss: 4.897519829683006e-05
step: 90, loss: 0.0005158839630894363
step: 100, loss: 0.004604771733283997
step: 110, loss: 8.270446414826438e-05
step: 120, loss: 3.0542702006641775e-05
step: 130, loss: 2.7274583771941252e-05
step: 140, loss: 0.00010311775986338034
step: 150, loss: 4.082830491825007e-05
step: 160, loss: 4.401059413794428e-05
step: 170, loss: 0.002458004280924797
step: 180, loss: 3.827177351922728e-05
step: 190, loss: 8.19078559288755e-05
step: 200, loss: 8.79556464497e-05
step: 210, loss: 2.8005853891954757e-05
step: 220, loss: 2.3450167645933107e-05
step: 230, loss: 1.976972089323681e-05
step: 240, loss: 4.067316695000045e-05
step: 250, loss: 2.6091322069987655e-05
step: 260, loss: 2.631084862514399e-05
step: 270, loss: 2.4757395294727758e-05
step: 280, loss: 4.316139529692009e-05
step: 290, loss: 7.615180948050693e-05
step: 300, loss: 6.591167039005086e-05
step: 310, loss: 0.0018230135319754481
step: 320, loss: 0.0076302820816636086
step: 330, loss: 0.00538630411028862
step: 340, loss: 3.187589391018264e-05
step: 350, loss: 3.062816176679917e-05
step: 360, loss: 0.02947108820080757
step: 370, loss: 3.518371158861555e-05
step: 380, loss: 4.580415043164976e-05

```

```

step: 390, loss: 3.669657598948106e-05
step: 400, loss: 2.6422667360748164e-05
step: 410, loss: 0.00012426861212588847
step: 420, loss: 3.578529504011385e-05
step: 430, loss: 0.007159748114645481
step: 440, loss: 7.798551087034866e-05
step: 450, loss: 6.180829223012552e-05
step: 460, loss: 0.0005916255759075284
step: 470, loss: 8.415251795668155e-05
step: 480, loss: 1.9617084035417065e-05
step: 490, loss: 5.03226910950616e-05
step: 500, loss: 1.9818131477222778e-05
step: 510, loss: 0.00018146986258216202
step: 520, loss: 0.008449992164969444
step: 530, loss: 3.0143117328407243e-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=0.9171428571428571, f1=0.914339801230478,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013865585788153112
step: 10, loss: 2.1092131646582857e-05
step: 20, loss: 6.898032006574795e-05
step: 30, loss: 2.5133938834187575e-05
step: 40, loss: 6.14163582213223e-05
step: 50, loss: 0.0001560638629598543
step: 60, loss: 4.204561628284864e-05
step: 70, loss: 0.0020134865771979094
step: 80, loss: 0.0014126162277534604
step: 90, loss: 0.00010403302439954132
step: 100, loss: 3.003950223501306e-05
step: 110, loss: 2.470177241775673e-05
step: 120, loss: 2.3192969820229337e-05
step: 130, loss: 0.0008930280455388129
step: 140, loss: 3.567373278201558e-05
step: 150, loss: 4.797432120540179e-05
step: 160, loss: 0.00010297936387360096
step: 170, loss: 5.289493128657341e-05
step: 180, loss: 2.7026135285268538e-05
step: 190, loss: 0.0001173585987999104
step: 200, loss: 4.991512469132431e-05
step: 210, loss: 2.3785507437423803e-05
step: 220, loss: 3.054941407754086e-05
step: 230, loss: 0.00012950049131177366
step: 240, loss: 5.028907617088407e-05

```

```

step: 250, loss: 4.476170215639286e-05
step: 260, loss: 0.00011719614849425852
step: 270, loss: 8.364575478481129e-05
step: 280, loss: 3.555903094820678e-05
step: 290, loss: 5.905404395889491e-05
step: 300, loss: 0.0093693183735013
step: 310, loss: 3.5764915082836524e-05
step: 320, loss: 8.659241575514898e-05
step: 330, loss: 5.5020893341861665e-05
step: 340, loss: 5.078493268229067e-05
step: 350, loss: 4.119370350963436e-05
step: 360, loss: 0.0005837864009663463
step: 370, loss: 0.00010801666940096766
step: 380, loss: 0.0006235428736545146
step: 390, loss: 0.02498939074575901
step: 400, loss: 0.00010978255158988759
step: 410, loss: 1.8953915059682913e-05
step: 420, loss: 2.3669837901252322e-05
step: 430, loss: 3.794037183979526e-05
step: 440, loss: 9.274420881411061e-05
step: 450, loss: 4.5257300371304154e-05
step: 460, loss: 2.5599380023777485e-05
step: 470, loss: 2.8757338441209868e-05
step: 480, loss: 2.4846875021466985e-05
step: 490, loss: 2.80873391602654e-05
step: 500, loss: 1.9222146875108592e-05
step: 510, loss: 0.0003137631865683943
step: 520, loss: 3.147298048133962e-05
step: 530, loss: 0.0009378571412526071
/usr/local/lib/python3.7/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=0.9211267605633803, f1=0.9206349206349207,
best_f1=0.9161290322580645
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019243224523961544
step: 10, loss: 2.530510028009303e-05
step: 20, loss: 3.2354470022255555e-05
step: 30, loss: 0.00010604789713397622
step: 40, loss: 8.114125375868753e-05
step: 50, loss: 0.001059560920111835
step: 60, loss: 1.923690433613956e-05
step: 70, loss: 0.0002254409482702613
step: 80, loss: 2.4258422854472883e-05
step: 90, loss: 0.0007244282751344144
step: 100, loss: 6.953974661882967e-05

```

```
step: 110, loss: 0.0005227940273471177
step: 120, loss: 8.53989549796097e-05
step: 130, loss: 0.0036613321863114834
step: 140, loss: 6.251762533793226e-05
step: 150, loss: 3.970854959334247e-05
step: 160, loss: 8.92361713340506e-05
step: 170, loss: 4.1762406908674166e-05
step: 180, loss: 1.7210548321600072e-05
step: 190, loss: 4.116906347917393e-05
step: 200, loss: 2.4548977307858877e-05
step: 210, loss: 8.927725139074028e-05
step: 220, loss: 0.00012316840002313256
step: 230, loss: 8.932689524954185e-05
step: 240, loss: 2.0902261894661933e-05
step: 250, loss: 5.516616874956526e-05
step: 260, loss: 0.00013078711344860494
step: 270, loss: 7.12586406734772e-05
step: 280, loss: 3.7814737879671156e-05
step: 290, loss: 0.04060002788901329
step: 300, loss: 7.642189302714542e-05
step: 310, loss: 0.00016089262499008328
step: 320, loss: 1.719936881272588e-05
step: 330, loss: 3.534551433403976e-05
step: 340, loss: 2.3837581466068514e-05
step: 350, loss: 2.185938683396671e-05
step: 360, loss: 0.0007592376205138862
step: 370, loss: 1.8246108083985746e-05
step: 380, loss: 6.406955071724951e-05
step: 390, loss: 3.735821883310564e-05
step: 400, loss: 2.7487893021316268e-05
step: 410, loss: 0.00026385675300844014
step: 420, loss: 2.9235150577733293e-05
step: 430, loss: 0.00012142696505179629
step: 440, loss: 4.559612716548145e-05
step: 450, loss: 0.00010717218538047746
step: 460, loss: 3.1186042178887874e-05
step: 470, loss: 0.0002379176439717412
step: 480, loss: 1.487120425736066e-05
step: 490, loss: 2.4228540496551432e-05
step: 500, loss: 2.1848371034138836e-05
step: 510, loss: 1.966915442608297e-05
step: 520, loss: 2.6780175176099874e-05
step: 530, loss: 0.0001834465510910377
/usr/local/lib/python3.7/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=0.9230055658627088, f1=0.9207373271889401,
best_f1=0.9161290322580645
```

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

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

```
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:24, 232.92it/s]
load_f1 = 0.9173138419369575
real_f1 = 0.9193621867881548
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, 244.29it/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: 0.5615615844726562
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 1: dev_f1=0.3448275862068965, f1=0.375, best_f1=0.375
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5297736525535583
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.43902439024390244, f1=0.43902439024390244,
best_f1=0.43902439024390244
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.5013142228126526
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5294117647058824, f1=0.43750000000000006,
best_f1=0.43750000000000006
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.29447659850120544
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5806451612903226, f1=0.4615384615384615,
best_f1=0.4615384615384615
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2740180492401123
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7142857142857143, f1=0.4827586206896552,
best_f1=0.4827586206896552
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.4226515293121338
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7200000000000001, f1=0.5833333333333334,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.3123817443847656
/usr/local/lib/python3.7/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=0.6285714285714286, f1=0.4285714285714286,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17127671837806702
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7407407407407408, f1=0.5833333333333334,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00497622461989522
/usr/local/lib/python3.7/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=0.7407407407407408, f1=0.5384615384615384,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06995583325624466
/usr/local/lib/python3.7/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=0.7142857142857143, f1=0.6153846153846153,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04421799257397652
/usr/local/lib/python3.7/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=0.6666666666666666, f1=0.6000000000000001,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01564813405275345
/usr/local/lib/python3.7/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=0.6666666666666667, f1=0.5294117647058824,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.01747032441198826
/usr/local/lib/python3.7/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=0.6666666666666667, f1=0.5625000000000001,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025611910969018936
/usr/local/lib/python3.7/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=0.6875000000000001, f1=0.5625000000000001,
best_f1=0.5833333333333334
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.018707094714045525
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 15: dev_f1=0.6875000000000001, f1=0.5625000000000001,
best_f1=0.5833333333333334
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, 123123.12it/s]
/usr/local/lib/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 = 0.689655172413793
real_f1 = 0.689655172413793
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, 241.05it/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: 0.5912836194038391
step: 10, loss: 0.6436368823051453
step: 20, loss: 0.4494129717350006
step: 30, loss: 0.25450095534324646
step: 40, loss: 0.2803192138671875
step: 50, loss: 0.05536786839365959
step: 60, loss: 0.07025075703859329
step: 70, loss: 0.35598334670066833
step: 80, loss: 0.09885349124670029
step: 90, loss: 0.14349022507667542
step: 100, loss: 0.042209792882204056
step: 110, loss: 0.1279958337545395
step: 120, loss: 0.013580985367298126
step: 130, loss: 0.008160744793713093
step: 140, loss: 0.013018233701586723
step: 150, loss: 0.03971901535987854
step: 160, loss: 0.011939644813537598
step: 170, loss: 0.045513588935136795
step: 180, loss: 0.1468939483165741
step: 190, loss: 0.0329338014125824
step: 200, loss: 0.03864526003599167
step: 210, loss: 0.05485939979553223
step: 220, loss: 0.01706445962190628
step: 230, loss: 0.04030580446124077
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9808342728297633, f1=0.9709821428571428,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009492904879152775
step: 10, loss: 0.010450083762407303
step: 20, loss: 0.1378997266292572
step: 30, loss: 0.18254120647907257
step: 40, loss: 0.016633134335279465
step: 50, loss: 0.014288675040006638
step: 60, loss: 0.0026654002722352743
step: 70, loss: 0.18263214826583862
step: 80, loss: 0.0030327499844133854
step: 90, loss: 0.025571485981345177
step: 100, loss: 0.01426483504474163
step: 110, loss: 0.03190067037940025
step: 120, loss: 0.027163799852132797
step: 130, loss: 0.01883687824010849
step: 140, loss: 0.009128754958510399
step: 150, loss: 0.051172301173210144
step: 160, loss: 0.004097449127584696
step: 170, loss: 0.003567297011613846
step: 180, loss: 0.0013276553945615888
step: 190, loss: 0.04449012875556946
step: 200, loss: 0.0019929339177906513
step: 210, loss: 0.005313188303261995
step: 220, loss: 0.09516651928424835
step: 230, loss: 0.02029440551996231
/usr/local/lib/python3.7/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=0.9765886287625419, f1=0.974472807991121,
best_f1=0.9709821428571428
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004059543367475271
step: 10, loss: 0.031813278794288635
step: 20, loss: 0.001109903329052031
step: 30, loss: 0.06427969038486481
step: 40, loss: 0.2138177454471588
step: 50, loss: 0.1278141736984253
step: 60, loss: 0.12090211361646652
step: 70, loss: 0.012329182587563992
step: 80, loss: 0.029261935502290726
step: 90, loss: 0.022010745480656624
step: 100, loss: 0.014927313663065434
step: 110, loss: 0.0006138274911791086
step: 120, loss: 0.020764341577887535
step: 130, loss: 0.0007095561595633626

```

```

step: 140, loss: 0.01250176876783371
step: 150, loss: 0.001024774624966085
step: 160, loss: 0.0058945827186107635
step: 170, loss: 0.013108275830745697
step: 180, loss: 0.018358562141656876
step: 190, loss: 0.05605647340416908
step: 200, loss: 0.03110646829009056
step: 210, loss: 0.0009866626933217049
step: 220, loss: 0.0007571495370939374
step: 230, loss: 0.0007868582033552229
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9843400447427293, f1=0.9690265486725664,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007826826767995954
step: 10, loss: 0.00038824439980089664
step: 20, loss: 0.0035975913051515818
step: 30, loss: 0.000494031875859946
step: 40, loss: 0.0012181713245809078
step: 50, loss: 0.0012680926593020558
step: 60, loss: 0.001146849012002349
step: 70, loss: 0.005112092941999435
step: 80, loss: 0.000657114724162966
step: 90, loss: 0.0064938426949083805
step: 100, loss: 0.0012669179122895002
step: 110, loss: 0.0071785179898142815
step: 120, loss: 0.018007125705480576
step: 130, loss: 0.001614624634385109
step: 140, loss: 0.00022864491620566696
step: 150, loss: 0.29455864429473877
step: 160, loss: 0.0007761695887893438
step: 170, loss: 0.06484705954790115
step: 180, loss: 0.000733162509277463
step: 190, loss: 0.019812917336821556
step: 200, loss: 0.0009241694933734834
step: 210, loss: 0.02187633514404297
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 220, loss: 0.0004444965743459761
step: 230, loss: 0.003322875825688243
/usr/local/lib/python3.7/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=0.9730337078651685, f1=0.9765886287625419,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006526106153614819
step: 10, loss: 0.0007505778921768069
step: 20, loss: 0.002874403027817607
step: 30, loss: 0.002404535422101617
step: 40, loss: 0.0029779451433569193
step: 50, loss: 0.0004954394535161555
step: 60, loss: 0.14960138499736786
step: 70, loss: 0.0005184091860428452
step: 80, loss: 0.0004447514656931162
step: 90, loss: 0.00035687899799086154
step: 100, loss: 0.0007942302036099136
step: 110, loss: 0.0002719898766372353
step: 120, loss: 0.0019382343161851168
step: 130, loss: 0.00478206854313612
step: 140, loss: 0.008591129444539547
step: 150, loss: 0.026790466159582138
step: 160, loss: 0.0005842240061610937
step: 170, loss: 0.00230504316277802
step: 180, loss: 0.0022164650727063417
step: 190, loss: 0.12431565672159195
step: 200, loss: 0.002274681581184268
step: 210, loss: 0.0013904864899814129
step: 220, loss: 0.0008734457660466433
step: 230, loss: 0.0007714052917435765
/usr/local/lib/python3.7/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=0.9853768278965129,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001113216276280582
step: 10, loss: 0.001316238776780665
step: 20, loss: 0.00048249319661408663
step: 30, loss: 0.0018970059463754296
step: 40, loss: 0.0020582550205290318
step: 50, loss: 0.0010689400369301438
step: 60, loss: 0.0002892140473704785
step: 70, loss: 0.02720535919070244
step: 80, loss: 0.0025691490154713392

```

```

step: 90, loss: 0.011655852198600769
step: 100, loss: 0.0015656236791983247
step: 110, loss: 0.00026328343665227294
step: 120, loss: 0.0003592222055885941
step: 130, loss: 0.0003426947514526546
step: 140, loss: 0.00010641995322657749
step: 150, loss: 0.010855337604880333
step: 160, loss: 0.0424010343849659
step: 170, loss: 0.010654699988663197
step: 180, loss: 0.002277534455060959
step: 190, loss: 0.03642309084534645
step: 200, loss: 0.00019589520525187254
step: 210, loss: 0.006579171866178513
step: 220, loss: 0.00024836030206643045
step: 230, loss: 0.007187440991401672
/usr/local/lib/python3.7/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=0.984304932735426, f1=0.9810479375696767,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001482901512645185
step: 10, loss: 0.00014107360038906336
step: 20, loss: 0.0016257783863693476
step: 30, loss: 0.00011860084487125278
step: 40, loss: 0.0002066227898467332
step: 50, loss: 0.00014287805242929608
step: 60, loss: 0.0012920330045744777
step: 70, loss: 0.00019354936375748366
step: 80, loss: 0.000285964400973171
step: 90, loss: 0.00023740832693874836
step: 100, loss: 0.0001783588668331504
step: 110, loss: 0.0037732466589659452
step: 120, loss: 6.422910519177094e-05
step: 130, loss: 0.0013329032808542252
step: 140, loss: 0.00018848368199542165
step: 150, loss: 0.0017741440096870065
step: 160, loss: 0.01902649737894535
step: 170, loss: 0.001775332959368825
step: 180, loss: 0.0006655720062553883
step: 190, loss: 0.015081035904586315
step: 200, loss: 0.017793871462345123
step: 210, loss: 0.00016178310033865273
step: 220, loss: 0.0008959562983363867
step: 230, loss: 0.0003360004920978099
/usr/local/lib/python3.7/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=0.9819004524886877, f1=0.9796839729119639,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00014973821816965938
step: 10, loss: 0.004486340098083019
step: 20, loss: 0.00021602939523290843
step: 30, loss: 0.0002866130671463907
step: 40, loss: 0.0005872789770364761
step: 50, loss: 0.0004134218906983733
step: 60, loss: 0.0003740238316822797
step: 70, loss: 0.00045882558333687484
step: 80, loss: 0.00028476782608777285
step: 90, loss: 0.0019301505526527762
step: 100, loss: 0.0005723671056330204
step: 110, loss: 0.00039842340629547834
step: 120, loss: 0.1413656771183014
step: 130, loss: 0.00013586375280283391
step: 140, loss: 0.0003238037752453238
step: 150, loss: 0.00022672228806186467
step: 160, loss: 0.0001800040336092934
step: 170, loss: 0.00015165418153628707
step: 180, loss: 0.00011064508726121858
step: 190, loss: 0.02164008840918541
step: 200, loss: 6.667581328656524e-05
step: 210, loss: 0.005621002987027168
step: 220, loss: 0.00040964363142848015
step: 230, loss: 9.638178744353354e-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=0.9808342728297633, f1=0.9842696629213483,
best_f1=0.9690265486725664
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.468300075037405e-05
step: 10, loss: 0.0017529526958242059
step: 20, loss: 0.003282067831605673
step: 30, loss: 0.00011082662240369245
step: 40, loss: 0.010249382816255093
step: 50, loss: 5.8719080698210746e-05
step: 60, loss: 0.0064128851518034935
step: 70, loss: 0.00020785628294106573
step: 80, loss: 0.00015208873082883656

```

```

step: 90, loss: 0.0001736625999910757
step: 100, loss: 0.0002860598615370691
step: 110, loss: 0.0012767110019922256
step: 120, loss: 4.792617983184755e-05
step: 130, loss: 4.374338095658459e-05
step: 140, loss: 0.00015597573656123132
step: 150, loss: 0.0010692719370126724
step: 160, loss: 6.354537617880851e-05
step: 170, loss: 0.0002506821183487773
step: 180, loss: 0.0001390523830195889
step: 190, loss: 0.00016340425645466894
step: 200, loss: 0.00017770972044672817
step: 210, loss: 0.07228420674800873
step: 220, loss: 0.0025755565147846937
step: 230, loss: 0.00098368467297405
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
UserWarning: Please also save or load the state of the optimizer when saving or
loading the scheduler.
    warnings.warn(SAVE_STATE_WARNING, UserWarning)
epoch 9: dev_f1=0.9864559819413092, f1=0.9831649831649831,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001715741236694157
step: 10, loss: 0.0011923820711672306
step: 20, loss: 0.0002296712773386389
step: 30, loss: 0.0002112572838086635
step: 40, loss: 0.011805467307567596
step: 50, loss: 0.0010750576620921493
step: 60, loss: 0.013978522270917892
step: 70, loss: 0.00028450990794226527
step: 80, loss: 0.00013330973160918802
step: 90, loss: 0.00015111682296264917
step: 100, loss: 0.0015833306824788451
step: 110, loss: 8.735875599086285e-05
step: 120, loss: 0.00015217128384392709
step: 130, loss: 8.345385867869481e-05
step: 140, loss: 0.00016340722504537553
step: 150, loss: 0.0019067659741267562
step: 160, loss: 6.118199235061184e-05
step: 170, loss: 3.552684574970044e-05
step: 180, loss: 0.00043602974619716406
step: 190, loss: 0.0020916673820465803
step: 200, loss: 5.2157931349938735e-05

```

```

step: 210, loss: 4.278585038264282e-05
step: 220, loss: 8.111945498967543e-05
step: 230, loss: 0.00018982998153660446
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.9865168539325843,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0005098332185298204
step: 10, loss: 6.167472747620195e-05
step: 20, loss: 9.891140507534146e-05
step: 30, loss: 9.65890649240464e-05
step: 40, loss: 0.0009112178231589496
step: 50, loss: 8.384117245441303e-05
step: 60, loss: 0.00017895438941195607
step: 70, loss: 6.67557178530842e-05
step: 80, loss: 0.04097673296928406
step: 90, loss: 0.028731804341077805
step: 100, loss: 5.966609751340002e-05
step: 110, loss: 9.232599404640496e-05
step: 120, loss: 0.00010250679770251736
step: 130, loss: 0.0002869335003197193
step: 140, loss: 0.00010223184654023498
step: 150, loss: 0.005001886282116175
step: 160, loss: 8.75632103998214e-05
step: 170, loss: 0.0007410160033032298
step: 180, loss: 0.0012955247657373548
step: 190, loss: 2.8386077246977948e-05
step: 200, loss: 0.0001418653700966388
step: 210, loss: 0.00018661361536942422
step: 220, loss: 4.070425347890705e-05
step: 230, loss: 0.00015194251318462193
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
epoch 11: dev_f1=0.9820627802690582, f1=0.984304932735426,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 4.379961683298461e-05
step: 10, loss: 6.120956095401198e-05
step: 20, loss: 4.8670644900994375e-05
step: 30, loss: 0.00031529649277217686
step: 40, loss: 5.863955448148772e-05

```

```

step: 50, loss: 0.0001070927974069491
step: 60, loss: 7.85038573667407e-05
step: 70, loss: 5.207861977396533e-05
step: 80, loss: 4.676470052800141e-05
step: 90, loss: 6.359694816637784e-05
step: 100, loss: 9.192447760142386e-05
step: 110, loss: 0.00019542794325388968
step: 120, loss: 7.904202357167378e-05
step: 130, loss: 0.0002962719008792192
step: 140, loss: 0.00012872045044787228
step: 150, loss: 6.77174684824422e-05
step: 160, loss: 6.470434891525656e-05
step: 170, loss: 0.00011033289774786681
step: 180, loss: 0.00019550370052456856
step: 190, loss: 0.0006971173570491374
step: 200, loss: 0.0003019634750671685
step: 210, loss: 0.0006922603934071958
step: 220, loss: 0.000360679579898715
step: 230, loss: 0.04525629058480263
/usr/local/lib/python3.7/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=0.9808342728297633, f1=0.984304932735426,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0007590592140331864
step: 10, loss: 0.000248692580498755
step: 20, loss: 0.0002583207096904516
step: 30, loss: 0.00015497690765187144
step: 40, loss: 7.97479588072747e-05
step: 50, loss: 4.1349085222464055e-05
step: 60, loss: 4.684122541220859e-05
step: 70, loss: 5.802487430628389e-05
step: 80, loss: 0.00010451512935105711
step: 90, loss: 0.000550613971427083
step: 100, loss: 6.178844341775402e-05
step: 110, loss: 0.0016755074029788375
step: 120, loss: 9.320996468886733e-05
step: 130, loss: 7.221900159493089e-05
step: 140, loss: 7.3314702603966e-05
step: 150, loss: 6.391266651917249e-05
step: 160, loss: 0.00010907967953244224
step: 170, loss: 9.250863513443619e-05
step: 180, loss: 6.576562736881897e-05
step: 190, loss: 6.40438447589986e-05
step: 200, loss: 0.00010067258699564263

```

```

step: 210, loss: 2.39563323702896e-05
step: 220, loss: 0.0015348652377724648
step: 230, loss: 4.6673481847392395e-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=0.9808342728297633, f1=0.9854096520763187,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 6.142279016785324e-05
step: 10, loss: 0.0005020553362555802
step: 20, loss: 0.0002667384978849441
step: 30, loss: 0.0008615604601800442
step: 40, loss: 0.0007373114349320531
step: 50, loss: 7.327970524784178e-05
step: 60, loss: 0.00019137798517476767
step: 70, loss: 0.0003075944259762764
step: 80, loss: 4.783759868587367e-05
step: 90, loss: 0.00013629972818307579
step: 100, loss: 5.8208424889016896e-05
step: 110, loss: 4.677263132180087e-05
step: 120, loss: 4.873882426181808e-05
step: 130, loss: 6.244253017939627e-05
step: 140, loss: 9.646185208112001e-05
step: 150, loss: 5.4032858315622434e-05
step: 160, loss: 6.380406557582319e-05
step: 170, loss: 2.639295780682005e-05
step: 180, loss: 6.232778105186298e-05
step: 190, loss: 5.717920794268139e-05
step: 200, loss: 3.584245132515207e-05
step: 210, loss: 6.422327714972198e-05
step: 220, loss: 0.00010595342610031366
step: 230, loss: 0.00018721884407568723
/usr/local/lib/python3.7/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=0.9807909604519773, f1=0.984304932735426,
best_f1=0.9831649831649831
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.985803677001968e-05
step: 10, loss: 2.6124460418941453e-05
step: 20, loss: 7.736351108178496e-05
step: 30, loss: 4.088234345545061e-05
step: 40, loss: 8.077503298409283e-05

```

```

step: 50, loss: 0.00023775709269102663
step: 60, loss: 0.00010171208850806579
step: 70, loss: 5.038761446485296e-05
step: 80, loss: 8.487808372592553e-05
step: 90, loss: 6.576386658707634e-05
step: 100, loss: 3.665166877908632e-05
step: 110, loss: 3.3715648896759376e-05
step: 120, loss: 6.384847074514255e-05
step: 130, loss: 6.718731310684234e-05
step: 140, loss: 3.962125992984511e-05
step: 150, loss: 0.0024699154309928417
step: 160, loss: 0.002059583319351077
step: 170, loss: 3.5708097129827365e-05
step: 180, loss: 0.00034649495501071215
step: 190, loss: 5.1423266995698214e-05
step: 200, loss: 0.00012106999929528683
step: 210, loss: 6.606968236155808e-05
step: 220, loss: 0.002273876452818513
step: 230, loss: 5.732826321036555e-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=0.9808342728297633, f1=0.984304932735426,
best_f1=0.9831649831649831
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, 196.06it/s]
load_f1 = 0.9832402234636871
real_f1 = 0.983277591973244
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, 230.90it/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: 0.6432169675827026  
step: 10, loss: 0.5794656276702881  
step: 20, loss: 0.5686971545219421  
step: 30, loss: 0.2894785702228546  
step: 40, loss: 0.26723000407218933  
step: 50, loss: 0.39409443736076355  
step: 60, loss: 0.07866831123828888
```

step: 70, loss: 0.16858913004398346
step: 80, loss: 0.06388065963983536
step: 90, loss: 0.22042101621627808
step: 100, loss: 0.07504530251026154
step: 110, loss: 0.11860810220241547
step: 120, loss: 0.07462351769208908
step: 130, loss: 0.053010813891887665
step: 140, loss: 0.09560681134462357
step: 150, loss: 0.12046237289905548
step: 160, loss: 0.05177341774106026
step: 170, loss: 0.24975648522377014
step: 180, loss: 0.08558251708745956
step: 190, loss: 0.01805344596505165
step: 200, loss: 0.1721978783607483
step: 210, loss: 0.025900324806571007
step: 220, loss: 0.17499904334545135
step: 230, loss: 0.17509828507900238
step: 240, loss: 0.05129426717758179
step: 250, loss: 0.07659067213535309
step: 260, loss: 0.15420085191726685
step: 270, loss: 0.017204510048031807
step: 280, loss: 0.04848505184054375
step: 290, loss: 0.3285023272037506
step: 300, loss: 0.11035525798797607
step: 310, loss: 0.19164392352104187
step: 320, loss: 0.07271929830312729
step: 330, loss: 0.0873781219124794
step: 340, loss: 0.1564044952392578
step: 350, loss: 0.019945407286286354
step: 360, loss: 0.0987618938088417
step: 370, loss: 0.07613836973905563
step: 380, loss: 0.03377092257142067
step: 390, loss: 0.25155937671661377
step: 400, loss: 0.2865198254585266
step: 410, loss: 0.05861658602952957
step: 420, loss: 0.10153212398290634
step: 430, loss: 0.1574670374393463
step: 440, loss: 0.028653167188167572
step: 450, loss: 0.03249174728989601
step: 460, loss: 0.10807108134031296
step: 470, loss: 0.10928919911384583
step: 480, loss: 0.10203718394041061
step: 490, loss: 0.14547939598560333
step: 500, loss: 0.10380413383245468
step: 510, loss: 0.02645495906472206
step: 520, loss: 0.039297446608543396
step: 530, loss: 0.00389574165455997
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.908921933085502, f1=0.9076492537313433,
best_f1=0.9076492537313433
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12372715026140213
step: 10, loss: 0.06790507584810257
step: 20, loss: 0.08405815064907074
step: 30, loss: 0.03098459541797638
step: 40, loss: 0.08674251288175583
step: 50, loss: 0.14887656271457672
step: 60, loss: 0.02997262217104435
step: 70, loss: 0.08427724242210388
step: 80, loss: 0.03831005468964577
step: 90, loss: 0.021467776969075203
step: 100, loss: 0.05835474655032158
step: 110, loss: 0.14676812291145325
step: 120, loss: 0.0968737080693245
step: 130, loss: 0.09822472184896469
step: 140, loss: 0.0504692979156971
step: 150, loss: 0.0873098075389862
step: 160, loss: 0.021702472120523453
step: 170, loss: 0.12617754936218262
step: 180, loss: 0.04307136312127113
step: 190, loss: 0.061948660761117935
step: 200, loss: 0.0300495233386755
step: 210, loss: 0.06996268033981323
step: 220, loss: 0.13182173669338226
step: 230, loss: 0.016093363985419273
step: 240, loss: 0.10938452184200287
step: 250, loss: 0.09020396322011948
step: 260, loss: 0.011481172405183315
step: 270, loss: 0.16227972507476807
step: 280, loss: 0.02999839559197426
step: 290, loss: 0.0578247494995594
step: 300, loss: 0.17288850247859955
step: 310, loss: 0.010233774781227112
step: 320, loss: 0.10799362510442734
step: 330, loss: 0.11553701758384705
step: 340, loss: 0.022246431559324265
step: 350, loss: 0.0027778707444667816
step: 360, loss: 0.0849435105919838

```

```

step: 370, loss: 0.10982630401849747
step: 380, loss: 0.07202284038066864
step: 390, loss: 0.08721145987510681
step: 400, loss: 0.06855747848749161
step: 410, loss: 0.07379887253046036
step: 420, loss: 0.04697937145829201
step: 430, loss: 0.020284224301576614
step: 440, loss: 0.017918352037668228
step: 450, loss: 0.044734615832567215
step: 460, loss: 0.052820056676864624
step: 470, loss: 0.08989686518907547
step: 480, loss: 0.1787126064300537
step: 490, loss: 0.014756029471755028
step: 500, loss: 0.25569868087768555
step: 510, loss: 0.021564403548836708
step: 520, loss: 0.10473179817199707
step: 530, loss: 0.08317268639802933
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.917397323488694, f1=0.9124594719777674,
best_f1=0.9124594719777674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.21491973102092743
step: 10, loss: 0.08811965584754944
step: 20, loss: 0.11296552419662476
step: 30, loss: 0.3435492515563965
step: 40, loss: 0.03693331405520439
step: 50, loss: 0.030944766476750374
step: 60, loss: 0.13097211718559265
step: 70, loss: 0.03636755794286728
step: 80, loss: 0.011376826092600822
step: 90, loss: 0.010365281254053116
step: 100, loss: 0.07347522675991058
step: 110, loss: 0.04070301353931427
step: 120, loss: 0.011416873894631863
step: 130, loss: 0.0106012849137187
step: 140, loss: 0.007945485413074493
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 150, loss: 0.03563738986849785
step: 160, loss: 0.009678767062723637
step: 170, loss: 0.13403894007205963

```

```

step: 180, loss: 0.029295608401298523
step: 190, loss: 0.02699006162583828
step: 200, loss: 0.03233785927295685
step: 210, loss: 0.03896472603082657
step: 220, loss: 0.06872522830963135
step: 230, loss: 0.061165954917669296
step: 240, loss: 0.002930302172899246
step: 250, loss: 0.014165237545967102
step: 260, loss: 0.04662686213850975
step: 270, loss: 0.020530162379145622
step: 280, loss: 0.1493040919303894
step: 290, loss: 0.006884922739118338
step: 300, loss: 0.04535877704620361
step: 310, loss: 0.007923120632767677
step: 320, loss: 0.04219873994588852
step: 330, loss: 0.007602186873555183
step: 340, loss: 0.047734957188367844
step: 350, loss: 0.010542590171098709
step: 360, loss: 0.047138385474681854
step: 370, loss: 0.03142872825264931
step: 380, loss: 0.03950284421443939
step: 390, loss: 0.08419828861951828
step: 400, loss: 0.015115098096430302
step: 410, loss: 0.004173108842223883
step: 420, loss: 0.3515644669532776
step: 430, loss: 0.027394477277994156
step: 440, loss: 0.035865526646375656
step: 450, loss: 0.06232418864965439
step: 460, loss: 0.02362065576016903
step: 470, loss: 0.06603717803955078
step: 480, loss: 0.00512697221711278
step: 490, loss: 0.023833785206079483
step: 500, loss: 0.004552891943603754
step: 510, loss: 0.0071605509147048
step: 520, loss: 0.054668501019477844
step: 530, loss: 0.1638786345720291
/usr/local/lib/python3.7/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=0.9142857142857143, f1=0.915068493150685,
best_f1=0.9124594719777674
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0063229333609342575
step: 10, loss: 0.04357580095529556
step: 20, loss: 0.004959931131452322
step: 30, loss: 0.05596844479441643

```

step: 40, loss: 0.06611236184835434
step: 50, loss: 0.03200319781899452
step: 60, loss: 0.008341671898961067
step: 70, loss: 0.004574760794639587
step: 80, loss: 0.005781657062470913
step: 90, loss: 0.03557342663407326
step: 100, loss: 0.015723608434200287
step: 110, loss: 0.1024933010339737
step: 120, loss: 0.0016787794884294271
step: 130, loss: 0.013273424468934536
step: 140, loss: 0.006591208279132843
step: 150, loss: 0.02018212154507637
step: 160, loss: 0.004481967072933912
step: 170, loss: 0.010077516548335552
step: 180, loss: 0.023227954283356667
step: 190, loss: 0.041533760726451874
step: 200, loss: 0.00250938069075346
step: 210, loss: 0.15837661921977997
step: 220, loss: 0.020147236064076424
step: 230, loss: 0.10637642443180084
step: 240, loss: 0.02436026930809021
step: 250, loss: 0.17344465851783752
step: 260, loss: 0.0044840131886303425
step: 270, loss: 0.03124120645225048
step: 280, loss: 0.16565090417861938
step: 290, loss: 0.017541727051138878
step: 300, loss: 0.0051193260587751865
step: 310, loss: 0.00711792754009366
step: 320, loss: 0.014753562398254871
step: 330, loss: 0.012768425047397614
step: 340, loss: 0.0041920519433915615
step: 350, loss: 0.015385298058390617
step: 360, loss: 0.006404640153050423
step: 370, loss: 0.0025959163904190063
step: 380, loss: 0.00903854239732027
step: 390, loss: 0.00356141640804708
step: 400, loss: 0.0007490938296541572
step: 410, loss: 0.036410290747880936
step: 420, loss: 0.003811038099229336
step: 430, loss: 0.0664651170372963
step: 440, loss: 0.00946042500436306
step: 450, loss: 0.006267940159887075
step: 460, loss: 0.08807282149791718
step: 470, loss: 0.003630895633250475
step: 480, loss: 0.043030012398958206
step: 490, loss: 0.014034615829586983
step: 500, loss: 0.006909334100782871
step: 510, loss: 0.012908844277262688

```

step: 520, loss: 0.09783407300710678
step: 530, loss: 0.03270401060581207
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9193400549954172, f1=0.9178966789667896,
best_f1=0.9178966789667896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05320234224200249
step: 10, loss: 0.03313231095671654
step: 20, loss: 0.005363529082387686
step: 30, loss: 0.001266652368940413
step: 40, loss: 0.010541656985878944
step: 50, loss: 0.02974873036146164
step: 60, loss: 0.029580647125840187
step: 70, loss: 0.04263077303767204
step: 80, loss: 0.0006083748885430396
step: 90, loss: 0.03048071265220642
step: 100, loss: 0.17725932598114014
step: 110, loss: 0.01825079694390297
step: 120, loss: 0.00465376628562808
step: 130, loss: 0.0007920297794044018
step: 140, loss: 0.0256983432918787
step: 150, loss: 0.01082962192595005
step: 160, loss: 0.002842097310349345
step: 170, loss: 0.010495824739336967
step: 180, loss: 0.003239026525989175
step: 190, loss: 0.011058175936341286
step: 200, loss: 0.011724643409252167
step: 210, loss: 0.06432536989450455
step: 220, loss: 0.0016907139215618372
step: 230, loss: 0.005815912503749132
step: 240, loss: 0.26368096470832825
step: 250, loss: 0.0020668315701186657
step: 260, loss: 0.001958453329280019
step: 270, loss: 0.0010813859989866614
step: 280, loss: 0.017363782972097397
step: 290, loss: 0.11472629010677338
step: 300, loss: 0.03555500507354736
step: 310, loss: 0.0016295072855427861
step: 320, loss: 0.021979423239827156
step: 330, loss: 0.10333732515573502

```



```
step: 340, loss: 0.0030905436724424362
step: 350, loss: 0.006412309594452381
step: 360, loss: 0.02246265672147274
step: 370, loss: 0.055154308676719666
step: 380, loss: 0.0017654376570135355
step: 390, loss: 0.00031861712341196835
step: 400, loss: 0.03653624281287193
step: 410, loss: 0.005327066406607628
step: 420, loss: 0.0009492857498116791
step: 430, loss: 0.004287893418222666
step: 440, loss: 0.057711198925971985
step: 450, loss: 0.007742025889456272
step: 460, loss: 0.03097725659608841
step: 470, loss: 0.008831956423819065
step: 480, loss: 0.009001576341688633
step: 490, loss: 0.0008534133667126298
step: 500, loss: 0.13545680046081543
step: 510, loss: 0.04456082731485367
step: 520, loss: 0.016704432666301727
step: 530, loss: 0.04602591693401337
/usr/local/lib/python3.7/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=0.914365933551708, f1=0.9121495327102803,
best_f1=0.9178966789667896
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.12769797444343567
step: 10, loss: 0.003997258376330137
step: 20, loss: 0.12135347723960876
step: 30, loss: 0.008720848709344864
step: 40, loss: 0.004984455183148384
step: 50, loss: 0.003978007007390261
step: 60, loss: 0.00034319827682338655
step: 70, loss: 0.002659222576767206
step: 80, loss: 0.03731967881321907
step: 90, loss: 0.0011126877507194877
step: 100, loss: 0.023373859003186226
step: 110, loss: 0.0004954396281391382
step: 120, loss: 0.01719234511256218
step: 130, loss: 0.0007505913381464779
step: 140, loss: 0.02619159407913685
step: 150, loss: 0.0016640652902424335
step: 160, loss: 0.020040346309542656
step: 170, loss: 0.017490878701210022
step: 180, loss: 0.0021748302970081568
step: 190, loss: 0.002928877482190728
```

```

step: 200, loss: 0.0018086194759234786
step: 210, loss: 0.025980006903409958
step: 220, loss: 0.0037411621306091547
step: 230, loss: 0.004229275975376368
step: 240, loss: 0.010301388800144196
step: 250, loss: 0.1579292267560959
step: 260, loss: 0.0011781086213886738
step: 270, loss: 0.05868798866868019
step: 280, loss: 0.0035838214680552483
step: 290, loss: 0.0019139035139232874
step: 300, loss: 0.0033390759490430355
step: 310, loss: 0.05265134572982788
step: 320, loss: 0.000401978351874277
step: 330, loss: 0.0009159045876003802
step: 340, loss: 0.1482921987771988
step: 350, loss: 0.004065051209181547
step: 360, loss: 0.02321203052997589
step: 370, loss: 0.03329147398471832
step: 380, loss: 0.02397562935948372
step: 390, loss: 0.015764113515615463
step: 400, loss: 0.03955477848649025
step: 410, loss: 0.029203860089182854
step: 420, loss: 0.006945027504116297
step: 430, loss: 0.0006569690885953605
step: 440, loss: 0.0016032017301768064
step: 450, loss: 0.005591391585767269
step: 460, loss: 0.0004577552026603371
step: 470, loss: 0.04750676825642586
step: 480, loss: 0.010568506084382534
step: 490, loss: 0.003995290491729975
step: 500, loss: 0.006236369721591473
step: 510, loss: 0.0009215911268256605
step: 520, loss: 0.005584217142313719
step: 530, loss: 0.008746778592467308
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9217230199166282, f1=0.9157070474435745,
best_f1=0.9157070474435745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008866103016771376
step: 10, loss: 0.03541533276438713

```

step: 20, loss: 0.0010589235462248325
step: 30, loss: 0.0009383137221448123
step: 40, loss: 0.007941137067973614
step: 50, loss: 0.018864214420318604
step: 60, loss: 0.008207503706216812
step: 70, loss: 0.00019126213737763464
step: 80, loss: 0.03145724907517433
step: 90, loss: 0.00024961470626294613
step: 100, loss: 0.0002034027420450002
step: 110, loss: 0.0022424482740461826
step: 120, loss: 0.13061772286891937
step: 130, loss: 0.004915175028145313
step: 140, loss: 0.003931502345949411
step: 150, loss: 0.0025763818994164467
step: 160, loss: 0.00016469540423713624
step: 170, loss: 0.0008645814377814531
step: 180, loss: 0.00029091144097037613
step: 190, loss: 0.0003357613168191165
step: 200, loss: 0.0007675864617340267
step: 210, loss: 0.002037388039752841
step: 220, loss: 0.0005657821893692017
step: 230, loss: 0.005670848768204451
step: 240, loss: 0.0036391434259712696
step: 250, loss: 0.03226909786462784
step: 260, loss: 0.00155393045861274
step: 270, loss: 0.02800709940493107
step: 280, loss: 0.004200147930532694
step: 290, loss: 0.0008345990208908916
step: 300, loss: 0.010614514350891113
step: 310, loss: 0.03656494617462158
step: 320, loss: 0.0006216850015334785
step: 330, loss: 0.0003608901461120695
step: 340, loss: 0.053733211010694504
step: 350, loss: 0.003547732951119542
step: 360, loss: 0.0010804266203194857
step: 370, loss: 0.0015452009392902255
step: 380, loss: 0.003904322860762477
step: 390, loss: 0.0015550954267382622
step: 400, loss: 0.005224541295319796
step: 410, loss: 0.0011256505968049169
step: 420, loss: 0.01394051406532526
step: 430, loss: 0.010264958254992962
step: 440, loss: 0.0011269109090790153
step: 450, loss: 0.04625053331255913
step: 460, loss: 0.019836213439702988
step: 470, loss: 0.07629935443401337
step: 480, loss: 0.06309514492750168
step: 490, loss: 0.0004997917567379773

```

step: 500, loss: 0.001338051282800734
step: 510, loss: 0.0008395395125262439
step: 520, loss: 0.04311465099453926
step: 530, loss: 0.01219343300908804
/usr/local/lib/python3.7/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=0.91728624535316, f1=0.9096804075961092,
best_f1=0.9157070474435745
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021484028548002243
step: 10, loss: 0.003282646182924509
step: 20, loss: 0.0005291769048199058
step: 30, loss: 0.0033440992701798677
step: 40, loss: 0.00623582536354661
step: 50, loss: 0.007723519578576088
step: 60, loss: 0.0003241439117118716
step: 70, loss: 0.00017739066970534623
step: 80, loss: 0.0005010910099372268
step: 90, loss: 0.001334032160229981
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0014582984149456024
step: 110, loss: 0.035171907395124435
step: 120, loss: 0.006415925920009613
step: 130, loss: 0.0015159603208303452
step: 140, loss: 0.01067729014903307
step: 150, loss: 0.0004892240976914763
step: 160, loss: 0.008320561610162258
step: 170, loss: 0.003964596427977085
step: 180, loss: 0.03204644098877907
step: 190, loss: 0.0032287738285958767
step: 200, loss: 0.002640477381646633
step: 210, loss: 0.002335937228053808
step: 220, loss: 0.00489013222977519
step: 230, loss: 0.0014675643760710955
step: 240, loss: 0.0006858453270979226
step: 250, loss: 0.004329627845436335
step: 260, loss: 0.06180361658334732
step: 270, loss: 0.0002338849735679105
step: 280, loss: 0.014675603248178959
step: 290, loss: 0.004660504404455423
step: 300, loss: 0.0022376207634806633
step: 310, loss: 0.04093267023563385
step: 320, loss: 0.0003302262048237026
step: 330, loss: 0.000507361488416791
step: 340, loss: 0.013088764622807503

```

```

step: 350, loss: 0.01198047585785389
step: 360, loss: 0.002794749103486538
step: 370, loss: 0.0006022340385243297
step: 380, loss: 0.02472503110766411
step: 390, loss: 0.0753994956612587
step: 400, loss: 0.06648391485214233
step: 410, loss: 0.0020583434961736202
step: 420, loss: 0.005995376501232386
step: 430, loss: 0.0111875981092453
step: 440, loss: 0.0021616367157548666
step: 450, loss: 0.0012473904062062502
step: 460, loss: 0.026968462392687798
step: 470, loss: 0.0007701216964051127
step: 480, loss: 0.002123443176969886
step: 490, loss: 0.005948388483375311
step: 500, loss: 0.005623057950288057
step: 510, loss: 0.0021883402951061726
step: 520, loss: 0.002976586576551199
step: 530, loss: 0.0007060857024043798
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9233627496516488, f1=0.9123783031988874,
best_f1=0.9123783031988874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006570807890966535
step: 10, loss: 0.0037700068205595016
step: 20, loss: 0.006749854423105717
step: 30, loss: 0.001152255805209279
step: 40, loss: 0.000977344112470746
step: 50, loss: 0.00014748437388334423
step: 60, loss: 7.818690937710926e-05
step: 70, loss: 0.00039197897422127426
step: 80, loss: 0.0064156921580433846
step: 90, loss: 0.00014216022100299597
step: 100, loss: 0.00012421095743775368
step: 110, loss: 0.00026075897039845586
step: 120, loss: 0.010750691406428814
step: 130, loss: 0.015329508110880852
step: 140, loss: 0.010573489591479301
step: 150, loss: 0.0009126697550527751
step: 160, loss: 0.0005626273923553526

```

```

step: 170, loss: 0.021081330254673958
step: 180, loss: 0.00023237497953232378
step: 190, loss: 0.004912355914711952
step: 200, loss: 0.0007061215583235025
step: 210, loss: 0.0003475769190117717
step: 220, loss: 0.01134338229894638
step: 230, loss: 0.00026853583403863013
step: 240, loss: 0.00945820938795805
step: 250, loss: 0.012632842175662518
step: 260, loss: 0.011957025155425072
step: 270, loss: 9.201082866638899e-05
step: 280, loss: 0.00029794833972118795
step: 290, loss: 0.011704761534929276
step: 300, loss: 0.00013564882101491094
step: 310, loss: 0.01381321344524622
step: 320, loss: 0.0003214622847735882
step: 330, loss: 0.0002718884206842631
step: 340, loss: 0.00010680103878257796
step: 350, loss: 0.0014255247078835964
step: 360, loss: 0.00014812014705967158
step: 370, loss: 0.0007048416882753372
step: 380, loss: 0.00016902580682653934
step: 390, loss: 0.00040158172487281263
step: 400, loss: 0.0002844981208909303
step: 410, loss: 0.0024590070825070143
step: 420, loss: 5.836074342369102e-05
step: 430, loss: 0.00017242766625713557
step: 440, loss: 0.0036670181434601545
step: 450, loss: 0.00022078544134274125
step: 460, loss: 0.0007721959846094251
step: 470, loss: 5.1276096201036125e-05
step: 480, loss: 0.0012391373747959733
step: 490, loss: 0.002271660603582859
step: 500, loss: 0.003783675841987133
step: 510, loss: 0.0016009424580261111
step: 520, loss: 0.0004539195797406137
step: 530, loss: 0.00012021960719721392
/usr/local/lib/python3.7/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=0.919815668202765, f1=0.909007773205304,
best_f1=0.9123783031988874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014526124577969313
step: 10, loss: 7.141355308704078e-05
step: 20, loss: 0.00012395459634717554

```

step: 30, loss: 5.716820669476874e-05
step: 40, loss: 0.014403571374714375
step: 50, loss: 0.0001546005514683202
step: 60, loss: 0.00016487171524204314
step: 70, loss: 0.0001779926533345133
step: 80, loss: 8.437834185315296e-05
step: 90, loss: 0.0004558057407848537
step: 100, loss: 0.00010274788655806333
step: 110, loss: 7.834178541088477e-05
step: 120, loss: 0.0007394906133413315
step: 130, loss: 0.00017049828602466732
step: 140, loss: 0.00019995101320091635
step: 150, loss: 0.0026101444382220507
step: 160, loss: 0.0006443235324695706
step: 170, loss: 0.0005669977399520576
step: 180, loss: 0.001215497963130474
step: 190, loss: 0.00030760240042582154
step: 200, loss: 0.002735296031460166
step: 210, loss: 0.001392025500535965
step: 220, loss: 0.003804252715781331
step: 230, loss: 0.00035513032344169915
step: 240, loss: 0.0021062667947262526
step: 250, loss: 0.00018456151883583516
step: 260, loss: 0.00024592041154392064
step: 270, loss: 0.00028763458249159157
step: 280, loss: 0.0007450900739058852
step: 290, loss: 0.0003389817720744759
step: 300, loss: 0.0001402539201080799
step: 310, loss: 0.0009115709690377116
step: 320, loss: 0.0036456366069614887
step: 330, loss: 0.0009176344610750675
step: 340, loss: 0.025405559688806534
step: 350, loss: 0.004322477150708437
step: 360, loss: 0.00024232096620835364
step: 370, loss: 0.0005230470560491085
step: 380, loss: 0.04495873674750328
step: 390, loss: 0.002632845425978303
step: 400, loss: 0.000547737639863044
step: 410, loss: 0.0012745721032842994
step: 420, loss: 0.004148493520915508
step: 430, loss: 0.003611496649682522
step: 440, loss: 0.005687338765710592
step: 450, loss: 0.00020500893879216164
step: 460, loss: 0.0009671714506112039
step: 470, loss: 0.001243944396264851
step: 480, loss: 0.0021665191743522882
step: 490, loss: 0.0013511236757040024
step: 500, loss: 0.015312404371798038

```

step: 510, loss: 0.000706340535543859
step: 520, loss: 0.0014085295842960477
step: 530, loss: 0.001508255722001195
/usr/local/lib/python3.7/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=0.9165103189493434, f1=0.9115456238361266,
best_f1=0.9123783031988874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006769376341253519
step: 10, loss: 0.007076040375977755
step: 20, loss: 0.0011369726853445172
step: 30, loss: 0.0016552356537431479
step: 40, loss: 0.0015684545505791903
step: 50, loss: 0.002679994096979499
step: 60, loss: 0.009369200095534325
step: 70, loss: 0.00015245706890709698
step: 80, loss: 0.00012929273361805826
step: 90, loss: 0.00022265853476710618
step: 100, loss: 0.0012252939632162452
step: 110, loss: 0.0002645125496201217
step: 120, loss: 0.002573452889919281
step: 130, loss: 0.00041808185051195323
step: 140, loss: 0.0024208223912864923
step: 150, loss: 0.0001882688229670748
step: 160, loss: 0.00013466094969771802
step: 170, loss: 0.0011592558585107327
step: 180, loss: 0.0005455777863971889
step: 190, loss: 0.0009196480386890471
step: 200, loss: 0.004846976138651371
step: 210, loss: 0.0016713475342839956
step: 220, loss: 0.0007246572058647871
step: 230, loss: 0.00014087455929256976
step: 240, loss: 0.00024051731452345848
step: 250, loss: 4.507824996835552e-05
step: 260, loss: 7.00874297763221e-05
step: 270, loss: 0.002262780675664544
step: 280, loss: 0.001268831198103726
step: 290, loss: 0.005287983454763889
step: 300, loss: 0.00012603425420820713
step: 310, loss: 0.0002156123227905482
step: 320, loss: 0.0006061005406081676
step: 330, loss: 0.0004350713570602238
step: 340, loss: 0.00010532131273066625
step: 350, loss: 0.0015183438081294298
step: 360, loss: 0.00019962441001553088

```



```

step: 370, loss: 0.0004090456059202552
step: 380, loss: 6.761395343346521e-05
step: 390, loss: 0.00017333627329207957
step: 400, loss: 0.00014783187361899763
step: 410, loss: 0.006086017936468124
step: 420, loss: 0.00019714314839802682
step: 430, loss: 6.73909526085481e-05
step: 440, loss: 0.0004028283292427659
step: 450, loss: 0.00023967173183336854
step: 460, loss: 0.0009789586765691638
step: 470, loss: 5.949044498265721e-05
step: 480, loss: 0.0004894300363957882
step: 490, loss: 0.00035989625030197203
step: 500, loss: 0.0016888212412595749
step: 510, loss: 0.0001776448916643858
step: 520, loss: 0.0005586285260505974
step: 530, loss: 4.7826069931034e-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=0.9201307800093415, f1=0.915270018621974,
best_f1=0.9123783031988874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001668635755777359
step: 10, loss: 9.601422789273784e-05
step: 20, loss: 0.00011722083581844345
step: 30, loss: 0.0002559900749474764
step: 40, loss: 0.00010899812332354486
step: 50, loss: 0.0037270488683134317
step: 60, loss: 0.0002242066984763369
step: 70, loss: 0.0010435784934088588
step: 80, loss: 0.00015406383317895234
step: 90, loss: 0.00015598170284647495
step: 100, loss: 0.0012697571655735373
step: 110, loss: 0.0060674757696688175
step: 120, loss: 0.0002429101150482893
step: 130, loss: 0.00013263044820632786
step: 140, loss: 5.328568659024313e-05
step: 150, loss: 4.959966099704616e-05
step: 160, loss: 6.878014391986653e-05
step: 170, loss: 0.0002351402654312551
step: 180, loss: 0.0003163288638461381
step: 190, loss: 0.00020670056983362883
step: 200, loss: 0.00010789995576487854
step: 210, loss: 0.001180600724183023
step: 220, loss: 8.17626787465997e-05

```

```

step: 230, loss: 0.00016491721908096224
step: 240, loss: 0.02004065178334713
step: 250, loss: 7.284760795300826e-05
step: 260, loss: 8.48357449285686e-05
step: 270, loss: 0.00011025104322470725
step: 280, loss: 0.0001815672731027007
step: 290, loss: 0.00040681182872503996
step: 300, loss: 0.0007071828586049378
step: 310, loss: 8.006983989616856e-05
step: 320, loss: 0.03172234818339348
step: 330, loss: 0.0013416141737252474
step: 340, loss: 0.0005816711927764118
step: 350, loss: 0.00014779796765651554
step: 360, loss: 0.02997656911611557
step: 370, loss: 0.00029991412884555757
step: 380, loss: 6.245491385925561e-05
step: 390, loss: 0.0008959879050962627
step: 400, loss: 0.000689466018229723
step: 410, loss: 0.0019390406087040901
step: 420, loss: 0.013543483801186085
step: 430, loss: 0.0007323969039134681
step: 440, loss: 0.0002677678712643683
step: 450, loss: 0.09057087451219559
step: 460, loss: 0.00018184339569415897
step: 470, loss: 0.0005838802317157388
step: 480, loss: 0.0037541273050010204
step: 490, loss: 0.0004530463193077594
step: 500, loss: 0.00011814069875981659
step: 510, loss: 0.0008226241334341466
step: 520, loss: 8.933801291277632e-05
step: 530, loss: 0.0004185196303296834
/usr/local/lib/python3.7/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=0.922077922077922, f1=0.9158964879852127,
best_f1=0.9123783031988874
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000635494536254555
step: 10, loss: 0.0005558306002058089
step: 20, loss: 0.0006653271266259253
step: 30, loss: 0.00011362097575329244
step: 40, loss: 7.71367849665694e-05
step: 50, loss: 0.0001423661014996469
step: 60, loss: 0.0001117146312026307
step: 70, loss: 0.00021158054005354643
step: 80, loss: 0.0006761016556993127

```

```
step: 90, loss: 0.005123398266732693
step: 100, loss: 0.00045350275468081236
step: 110, loss: 9.313018381362781e-05
step: 120, loss: 0.0006336967344395816
step: 130, loss: 8.268420788226649e-05
step: 140, loss: 0.00026145210722461343
step: 150, loss: 8.401733794016764e-05
step: 160, loss: 0.0001823462371248752
step: 170, loss: 0.00011808936687884852
step: 180, loss: 4.023152359877713e-05
step: 190, loss: 0.00016359976143576205
step: 200, loss: 5.586137922364287e-05
step: 210, loss: 0.0005271054105833173
step: 220, loss: 4.663785875891335e-05
step: 230, loss: 0.00022744160378351808
step: 240, loss: 0.0008395200711674988
step: 250, loss: 8.997876284411177e-05
step: 260, loss: 5.294408038025722e-05
step: 270, loss: 3.562020356184803e-05
step: 280, loss: 0.0006562484195455909
step: 290, loss: 0.007479868363589048
step: 300, loss: 0.00012616832100320607
step: 310, loss: 0.0009010815410874784
step: 320, loss: 0.000309580413158983
step: 330, loss: 0.0003290224994998425
step: 340, loss: 0.0003337975940667093
step: 350, loss: 7.566367276012897e-05
step: 360, loss: 4.2107807530555874e-05
step: 370, loss: 0.00034325855085626245
step: 380, loss: 9.540552855469286e-05
step: 390, loss: 0.0006570870755240321
step: 400, loss: 7.716954132774845e-05
step: 410, loss: 0.0008609858923591673
step: 420, loss: 6.481459422502667e-05
step: 430, loss: 0.0005121629219502211
step: 440, loss: 0.07560984790325165
step: 450, loss: 0.0011715992586687207
step: 460, loss: 0.0008879017550498247
step: 470, loss: 0.00011436655768193305
step: 480, loss: 5.5688622524030507e-05
step: 490, loss: 0.0006861565634608269
step: 500, loss: 6.882134039187804e-05
step: 510, loss: 0.00010412819392513484
step: 520, loss: 0.008197702467441559
step: 530, loss: 0.001788068562746048
/usr/local/lib/python3.7/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=0.9206496519721578, f1=0.9131238447319778,  
best_f1=0.9123783031988874  
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:  
An input tensor was not cuda.  
    warnings.warn("An input tensor was not cuda.")  
step: 0, loss: 5.203404725762084e-05  
step: 10, loss: 4.099066791241057e-05  
step: 20, loss: 0.0002334219025215134  
step: 30, loss: 5.1569371862569824e-05  
step: 40, loss: 0.00010371125972596928  
step: 50, loss: 0.00014137380640022457  
step: 60, loss: 4.458108378457837e-05  
step: 70, loss: 0.0004566086863633245  
step: 80, loss: 0.00011357384937582538  
step: 90, loss: 7.824871863704175e-05  
step: 100, loss: 0.0006471999222412705  
step: 110, loss: 6.441182631533593e-05  
step: 120, loss: 7.201216067187488e-05  
step: 130, loss: 0.000551077420823276  
step: 140, loss: 0.00011486210132716224  
step: 150, loss: 0.01001254003494978  
step: 160, loss: 0.0001550538290757686  
step: 170, loss: 8.660005551064387e-05  
step: 180, loss: 0.000693858484737575  
step: 190, loss: 0.0002414926275378093  
step: 200, loss: 0.0032146754674613476  
step: 210, loss: 5.199792212806642e-05  
step: 220, loss: 6.773033237550408e-05  
step: 230, loss: 0.000227677752263844  
step: 240, loss: 0.00015030261420179158  
step: 250, loss: 0.00025461838231422007  
step: 260, loss: 0.0014768311521038413  
step: 270, loss: 0.0009630474378354847  
step: 280, loss: 0.013576949946582317  
step: 290, loss: 0.00039297007606364787  
step: 300, loss: 0.0001411733974236995  
step: 310, loss: 0.0004454281006474048  
step: 320, loss: 0.0001259307173313573  
step: 330, loss: 0.0002775034517981112  
step: 340, loss: 0.00013378339644987136  
step: 350, loss: 0.0002469460596330464  
step: 360, loss: 0.0020715745631605387  
step: 370, loss: 6.215545727172866e-05  
step: 380, loss: 0.0025776776019483805  
step: 390, loss: 0.0026483305264264345  
step: 400, loss: 0.000712643435690552  
step: 410, loss: 0.0012287645367905498  
step: 420, loss: 8.840048394631594e-05
```

```

step: 430, loss: 0.01243277546018362
step: 440, loss: 0.0006866753683425486
step: 450, loss: 0.0004922955413348973
step: 460, loss: 8.177180279744789e-05
step: 470, loss: 7.203661516541615e-05
step: 480, loss: 7.298558921320364e-05
step: 490, loss: 0.0001625900185899809
step: 500, loss: 7.350657688220963e-05
step: 510, loss: 0.0002855872444342822
step: 520, loss: 0.00031426240457221866
step: 530, loss: 0.00012277478526812047
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9237209302325581, f1=0.917941585535466,
best_f1=0.917941585535466
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00010532478336244822
step: 10, loss: 0.00034633479663170874
step: 20, loss: 0.004244769457727671
step: 30, loss: 0.0011896772775799036
step: 40, loss: 0.062770776450634
step: 50, loss: 0.03267700597643852
step: 60, loss: 7.535697659477592e-05
step: 70, loss: 0.0008974058437161148
step: 80, loss: 0.0005324697121977806
step: 90, loss: 0.0001245082530658692
step: 100, loss: 0.00043774733785539865
step: 110, loss: 0.0002854544436559081
step: 120, loss: 0.00035022301017306745
step: 130, loss: 0.01297916378825903
step: 140, loss: 0.00021239528723526746
step: 150, loss: 0.00020546908490359783
step: 160, loss: 0.0006706533022224903
step: 170, loss: 0.003662377130240202
step: 180, loss: 0.00022868625819683075
step: 190, loss: 0.003206589026376605
step: 200, loss: 0.00015848140174057335
step: 210, loss: 0.0013659356627613306
step: 220, loss: 0.00021475725225172937
step: 230, loss: 0.0009188464027829468
step: 240, loss: 0.00013192091137170792

```

```

step: 250, loss: 7.947587437229231e-05
step: 260, loss: 0.0001990810123970732
step: 270, loss: 7.086074037943035e-05
step: 280, loss: 0.00012221535143908113
step: 290, loss: 5.598340430879034e-05
step: 300, loss: 7.618069503223523e-05
step: 310, loss: 0.0006058196304365993
step: 320, loss: 0.00031114002922549844
step: 330, loss: 9.128131932811812e-05
step: 340, loss: 0.00030869824695400894
step: 350, loss: 0.00024088901409413666
step: 360, loss: 0.00015273159078788012
step: 370, loss: 0.0001511833688709885
step: 380, loss: 0.00036148412618786097
step: 390, loss: 0.0008466944564133883
step: 400, loss: 0.0030352266039699316
step: 410, loss: 0.0005577703705057502
step: 420, loss: 0.00012029075878672302
step: 430, loss: 0.00027130270609632134
step: 440, loss: 0.0027798719238489866
step: 450, loss: 3.993383143097162e-05
step: 460, loss: 8.820155926514417e-05
step: 470, loss: 0.006896992679685354
step: 480, loss: 5.4895626817597076e-05
step: 490, loss: 0.00622172188013792
step: 500, loss: 7.773871038807556e-05
step: 510, loss: 0.00014765124069526792
step: 520, loss: 0.001406090217642486
step: 530, loss: 9.08069487195462e-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=0.921999065857076, f1=0.9191449814126395,
best_f1=0.917941585535466
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 [00:29, 193.43it/s]
load_f1 = 0.9191449814126395
real_f1 = 0.9190938511326862
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 [00:22, 191.26it/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 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: 0.561932384967804
step: 10, loss: 0.38196662068367004
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 20, loss: 0.37541916966438293
step: 30, loss: 0.35598787665367126
step: 40, loss: 0.17340539395809174
step: 50, loss: 0.45658838748931885
step: 60, loss: 0.3421810567378998
step: 70, loss: 0.2333104908466339
step: 80, loss: 0.1929652988910675
step: 90, loss: 0.44718387722969055
step: 100, loss: 0.5071896910667419
step: 110, loss: 0.2838827073574066
step: 120, loss: 0.2185257375240326
step: 130, loss: 0.2789377272129059
step: 140, loss: 0.31536340713500977
step: 150, loss: 0.31696629524230957
step: 160, loss: 0.27377820014953613
step: 170, loss: 0.3022902011871338
step: 180, loss: 0.0815606489777565
step: 190, loss: 0.28965145349502563
step: 200, loss: 0.29116880893707275
step: 210, loss: 0.20195452868938446
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5304518664047151, f1=0.5251396648044692,
best_f1=0.5251396648044692
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.14377640187740326
step: 10, loss: 0.2467467188835144
step: 20, loss: 0.16409209370613098
step: 30, loss: 0.32453668117523193
step: 40, loss: 0.21309612691402435
step: 50, loss: 0.12727226316928864
step: 60, loss: 0.3512710928916931
step: 70, loss: 0.14912433922290802
step: 80, loss: 0.2837107479572296
step: 90, loss: 0.11995261907577515

```

```

step: 100, loss: 0.07547318935394287
step: 110, loss: 0.17647330462932587
step: 120, loss: 0.18774248659610748
step: 130, loss: 0.04876979440450668
step: 140, loss: 0.22882550954818726
step: 150, loss: 0.2446465790271759
step: 160, loss: 0.21690145134925842
step: 170, loss: 0.14461781084537506
step: 180, loss: 0.2507364749908447
step: 190, loss: 0.2718774974346161
step: 200, loss: 0.07781342417001724
step: 210, loss: 0.17921504378318787
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5394321766561515, f1=0.5681818181818182,
best_f1=0.5681818181818182
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10968358814716339
step: 10, loss: 0.16632553935050964
step: 20, loss: 0.19120532274246216
step: 30, loss: 0.31611472368240356
step: 40, loss: 0.06593797355890274
step: 50, loss: 0.07002276182174683
step: 60, loss: 0.2627519965171814
step: 70, loss: 0.20173591375350952
step: 80, loss: 0.17694003880023956
step: 90, loss: 0.10377304255962372
step: 100, loss: 0.12604191899299622
step: 110, loss: 0.08811713010072708
step: 120, loss: 0.24525734782218933
step: 130, loss: 0.15299075841903687
step: 140, loss: 0.20891696214675903
step: 150, loss: 0.22955092787742615
step: 160, loss: 0.062040429562330246
step: 170, loss: 0.1881375014781952
step: 180, loss: 0.10262472182512283
step: 190, loss: 0.2185002714395523
step: 200, loss: 0.06665173172950745
step: 210, loss: 0.1326162964105606
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5613382899628253, f1=0.5813528336380257,
best_f1=0.5813528336380257
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.23238737881183624
step: 10, loss: 0.10036902874708176
step: 20, loss: 0.26490330696105957
step: 30, loss: 0.3516039550304413
step: 40, loss: 0.08425695449113846
step: 50, loss: 0.1862778514623642
step: 60, loss: 0.23209835588932037
step: 70, loss: 0.35358524322509766
step: 80, loss: 0.09305718541145325
step: 90, loss: 0.07219663262367249
step: 100, loss: 0.30651605129241943
step: 110, loss: 0.2245236188173294
step: 120, loss: 0.15707901120185852
step: 130, loss: 0.1405888795852661
step: 140, loss: 0.19224271178245544
step: 150, loss: 0.15436334908008575
step: 160, loss: 0.044449228793382645
step: 170, loss: 0.11533539742231369
step: 180, loss: 0.42306676506996155
step: 190, loss: 0.134471595287323
step: 200, loss: 0.16183386743068695
step: 210, loss: 0.27163609862327576
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.5720164609053499, f1=0.5778688524590164,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.10064142942428589
step: 10, loss: 0.06308963149785995
step: 20, loss: 0.35051581263542175
step: 30, loss: 0.06231073662638664

```

```

step: 40, loss: 0.16383615136146545
step: 50, loss: 0.16960714757442474
step: 60, loss: 0.1653696447610855
step: 70, loss: 0.22835950553417206
step: 80, loss: 0.07332412898540497
step: 90, loss: 0.10657387226819992
step: 100, loss: 0.009353180415928364
step: 110, loss: 0.3601433336734772
step: 120, loss: 0.06742251664400101
step: 130, loss: 0.14123821258544922
step: 140, loss: 0.13999442756175995
step: 150, loss: 0.07046673446893692
step: 160, loss: 0.17797496914863586
step: 170, loss: 0.0947832465171814
step: 180, loss: 0.11231426894664764
step: 190, loss: 0.08896876126527786
step: 200, loss: 0.05526620149612427
step: 210, loss: 0.0389108844101429
/usr/local/lib/python3.7/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=0.5458089668615984, f1=0.5632183908045977,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05238144472241402
step: 10, loss: 0.07270846515893936
step: 20, loss: 0.0771159827709198
step: 30, loss: 0.004129712004214525
step: 40, loss: 0.07813098281621933
step: 50, loss: 0.016708573326468468
step: 60, loss: 0.186320960521698
step: 70, loss: 0.01720019057393074
step: 80, loss: 0.07310400158166885
step: 90, loss: 0.22228090465068817
step: 100, loss: 0.012423837557435036
step: 110, loss: 0.047253429889678955
step: 120, loss: 0.0178973451256752
step: 130, loss: 0.05443910136818886
step: 140, loss: 0.027176018804311752
step: 150, loss: 0.030436711385846138
step: 160, loss: 0.012936200946569443
step: 170, loss: 0.1433609426021576
step: 180, loss: 0.019354302436113358
step: 190, loss: 0.22401508688926697
step: 200, loss: 0.009117942303419113
step: 210, loss: 0.047167591750621796

```

```

/usr/local/lib/python3.7/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=0.5378787878787878, f1=0.540952380952381,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05575336515903473
step: 10, loss: 0.03062710538506508
step: 20, loss: 0.02046326920390129
step: 30, loss: 0.141821026802063
step: 40, loss: 0.06264271587133408
step: 50, loss: 0.09894751012325287
step: 60, loss: 0.10437338799238205
step: 70, loss: 0.032828912138938904
step: 80, loss: 0.06011096015572548
step: 90, loss: 0.11172237247228622
step: 100, loss: 0.013100172393023968
step: 110, loss: 0.17753493785858154
step: 120, loss: 0.3318089544773102
step: 130, loss: 0.0778261125087738
step: 140, loss: 0.033866189420223236
step: 150, loss: 0.13273563981056213
step: 160, loss: 0.08648640662431717
step: 170, loss: 0.0255570225417614
step: 180, loss: 0.04886248707771301
step: 190, loss: 0.029400870203971863
step: 200, loss: 0.03392895311117172
step: 210, loss: 0.20769844949245453
/usr/local/lib/python3.7/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=0.5381526104417671, f1=0.511340206185567,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.028750427067279816
step: 10, loss: 0.15804320573806763
step: 20, loss: 0.02598012425005436
step: 30, loss: 0.06585128605365753
step: 40, loss: 0.06588435918092728
step: 50, loss: 0.10918564349412918
step: 60, loss: 0.13216708600521088
step: 70, loss: 0.11248582601547241
step: 80, loss: 0.08318179100751877
step: 90, loss: 0.06125306710600853

```

```

step: 100, loss: 0.12831173837184906
step: 110, loss: 0.08659306913614273
step: 120, loss: 0.004709347151219845
step: 130, loss: 0.02683163620531559
step: 140, loss: 0.16099798679351807
step: 150, loss: 0.12807050347328186
step: 160, loss: 0.05142957717180252
step: 170, loss: 0.016219468787312508
step: 180, loss: 0.10748087614774704
step: 190, loss: 0.06535530835390091
step: 200, loss: 0.0026355963200330734
step: 210, loss: 0.04204375296831131
/usr/local/lib/python3.7/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=0.5571142284569138, f1=0.5295315682281059,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02137613482773304
step: 10, loss: 0.14480091631412506
step: 20, loss: 0.008939669467508793
step: 30, loss: 0.13524125516414642
step: 40, loss: 0.05369008705019951
step: 50, loss: 0.05098594352602959
step: 60, loss: 0.12680085003376007
step: 70, loss: 0.11355870962142944
step: 80, loss: 0.01467288937419653
step: 90, loss: 0.0031887292861938477
step: 100, loss: 0.06404727697372437
step: 110, loss: 0.08183549344539642
step: 120, loss: 0.06538332253694534
step: 130, loss: 0.01797512359917164
step: 140, loss: 0.13163688778877258
step: 150, loss: 0.23596611618995667
step: 160, loss: 0.008830592036247253
step: 170, loss: 0.04893770441412926
step: 180, loss: 0.036039821803569794
step: 190, loss: 0.00711718900129199
step: 200, loss: 0.12601400911808014
step: 210, loss: 0.05112858861684799
/usr/local/lib/python3.7/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=0.5390946502057613, f1=0.5274261603375527,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:

```

```

An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.057605382055044174
step: 10, loss: 0.051808472722768784
step: 20, loss: 0.004706875886768103
step: 30, loss: 0.35790812969207764
step: 40, loss: 0.0074671427719295025
step: 50, loss: 0.005264155566692352
step: 60, loss: 0.12026776373386383
step: 70, loss: 0.03628760203719139
step: 80, loss: 0.06649953126907349
step: 90, loss: 0.06965833902359009
step: 100, loss: 0.012888294644653797
step: 110, loss: 0.00446836743503809
step: 120, loss: 0.0019691421184688807
step: 130, loss: 0.11453203856945038
step: 140, loss: 0.0032230087090283632
step: 150, loss: 0.028772303834557533
step: 160, loss: 0.02403487265110016
step: 170, loss: 0.0013290190836414695
step: 180, loss: 0.10206001996994019
step: 190, loss: 0.13138803839683533
step: 200, loss: 0.023523278534412384
step: 210, loss: 0.0836658626794815
/usr/local/lib/python3.7/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=0.532, f1=0.4968152866242039, best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0549805723130703
step: 10, loss: 0.10940007865428925
step: 20, loss: 0.08297421038150787
step: 30, loss: 0.00503826467320323
step: 40, loss: 0.012411331757903099
step: 50, loss: 0.14142844080924988
step: 60, loss: 0.05044317618012428
step: 70, loss: 0.027226148173213005
step: 80, loss: 0.14131534099578857
step: 90, loss: 0.02245078794658184
step: 100, loss: 0.0032529335003346205
step: 110, loss: 0.016533832997083664
step: 120, loss: 0.07380200177431107
step: 130, loss: 0.013233800418674946
step: 140, loss: 0.19027279317378998
step: 150, loss: 0.00043458197615109384
step: 160, loss: 0.002324250526726246

```

```

step: 170, loss: 0.05941487476229668
step: 180, loss: 0.004099602811038494
step: 190, loss: 0.027812723070383072
step: 200, loss: 0.07288870960474014
step: 210, loss: 0.0035410395357757807
/usr/local/lib/python3.7/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=0.5476673427991886, f1=0.526530612244898,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.04492562636733055
step: 10, loss: 0.0032380041666328907
step: 20, loss: 0.21624456346035004
step: 30, loss: 0.026094721630215645
step: 40, loss: 0.06622488796710968
step: 50, loss: 0.0036453185603022575
step: 60, loss: 0.010420555248856544
step: 70, loss: 0.002351746428757906
step: 80, loss: 0.04536993429064751
step: 90, loss: 0.00308625097386539
step: 100, loss: 0.009059763513505459
step: 110, loss: 0.002124052494764328
step: 120, loss: 0.018663981929421425
step: 130, loss: 0.02559361420571804
step: 140, loss: 0.07646898180246353
step: 150, loss: 0.016304928809404373
step: 160, loss: 0.0027268726844340563
step: 170, loss: 0.030342960730195045
step: 180, loss: 0.033858492970466614
step: 190, loss: 0.04353882744908333
step: 200, loss: 0.002245499985292554
step: 210, loss: 0.01419768575578928
/usr/local/lib/python3.7/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=0.5348314606741573, f1=0.5342465753424658,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.07868249714374542
step: 10, loss: 0.0003514767449814826
step: 20, loss: 0.012256370857357979
step: 30, loss: 0.06384357810020447
step: 40, loss: 0.08444519340991974

```



```

step: 50, loss: 0.008383489213883877
step: 60, loss: 0.017166731879115105
step: 70, loss: 0.03905710577964783
step: 80, loss: 0.02983442135155201
step: 90, loss: 0.0019172948086634278
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 100, loss: 0.0012057790299877524
step: 110, loss: 0.0031551369465887547
step: 120, loss: 0.0028723953291773796
step: 130, loss: 0.0075820498168468475
step: 140, loss: 0.0012145208893343806
step: 150, loss: 0.006989024579524994
step: 160, loss: 0.010570296086370945
step: 170, loss: 0.00286588235758245
step: 180, loss: 0.03306988254189491
step: 190, loss: 0.0010389761300757527
step: 200, loss: 0.011803828179836273
step: 210, loss: 0.03821675479412079
/usr/local/lib/python3.7/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=0.5272331154684096, f1=0.5067873303167422,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0034852023236453533
step: 10, loss: 0.033530380576848984
step: 20, loss: 0.001821668352931738
step: 30, loss: 0.0022065567318350077
step: 40, loss: 0.002641289494931698
step: 50, loss: 0.0497792512178421
step: 60, loss: 0.0015497613931074739
step: 70, loss: 0.0033847633749246597
step: 80, loss: 0.009441709145903587
step: 90, loss: 0.011456570588052273
step: 100, loss: 0.012943138368427753
step: 110, loss: 0.0012560974573716521
step: 120, loss: 0.051594335585832596
step: 130, loss: 0.15184339880943298
step: 140, loss: 0.02382732555270195
step: 150, loss: 0.002482305746525526
step: 160, loss: 0.009214404970407486
step: 170, loss: 0.0006162370555102825
step: 180, loss: 0.05037083849310875
step: 190, loss: 0.05044228583574295
step: 200, loss: 0.005049329251050949
step: 210, loss: 0.04710579290986061

```

```

/usr/local/lib/python3.7/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=0.5188284518828452, f1=0.524945770065076,
best_f1=0.5778688524590164
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0019311351934447885
step: 10, loss: 0.0011363789672031999
step: 20, loss: 0.01021649967879057
step: 30, loss: 0.19968253374099731
step: 40, loss: 0.035196613520383835
step: 50, loss: 0.02234460413455963
step: 60, loss: 0.013750078156590462
step: 70, loss: 0.000797578482888639
step: 80, loss: 0.01576909050345421
step: 90, loss: 0.007913421839475632
step: 100, loss: 0.2132837474346161
step: 110, loss: 0.0007412023842334747
step: 120, loss: 0.002454939531162381
step: 130, loss: 0.04487837851047516
step: 140, loss: 0.0011521333362907171
step: 150, loss: 0.0007193986675702035
step: 160, loss: 0.007811408489942551
step: 170, loss: 0.002057329285889864
step: 180, loss: 0.003477703081443906
step: 190, loss: 0.013397534377872944
step: 200, loss: 0.018091943114995956
step: 210, loss: 0.023180348798632622
/usr/local/lib/python3.7/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=0.5176991150442478, f1=0.5034642032332564,
best_f1=0.5778688524590164
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, 271.49it/s]

load_f1 = 0.592436974789916

real_f1 = 0.5889830508474577

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, 186.14it/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 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: 0.5842219591140747
step: 10, loss: 0.36847999691963196
step: 20, loss: 0.28939950466156006
step: 30, loss: 0.4428635537624359
step: 40, loss: 0.4337776303291321
step: 50, loss: 0.2991374731063843
step: 60, loss: 0.28736555576324463
step: 70, loss: 0.2878355085849762
step: 80, loss: 0.20732273161411285
step: 90, loss: 0.29357412457466125
step: 100, loss: 0.3222982585430145
step: 110, loss: 0.45145344734191895
step: 120, loss: 0.1955007165670395
step: 130, loss: 0.12506334483623505
step: 140, loss: 0.13794009387493134
step: 150, loss: 0.20704123377799988
step: 160, loss: 0.14964592456817627
step: 170, loss: 0.2752370536327362
step: 180, loss: 0.018978780135512352
step: 190, loss: 0.21886348724365234
```

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

```
warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6432160804020101, f1=0.6255924170616114,
best_f1=0.6255924170616114
```

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

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.24262726306915283
step: 10, loss: 0.08228439837694168
step: 20, loss: 0.18907266855239868
step: 30, loss: 0.15480133891105652
step: 40, loss: 0.22331440448760986
step: 50, loss: 0.4466925859451294
step: 60, loss: 0.40300533175468445
step: 70, loss: 0.3136136829853058
step: 80, loss: 0.22930724918842316
step: 90, loss: 0.22922837734222412
step: 100, loss: 0.05980762094259262
step: 110, loss: 0.13692079484462738
```

```

step: 120, loss: 0.26930996775627136
step: 130, loss: 0.11059590429067612
step: 140, loss: 0.09162820875644684
step: 150, loss: 0.16995376348495483
step: 160, loss: 0.04909989610314369
step: 170, loss: 0.1814032346010208
step: 180, loss: 0.16269414126873016
step: 190, loss: 0.05534593015909195
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.6666666666666666, f1=0.721763085399449,
best_f1=0.721763085399449
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.061973292380571365
step: 10, loss: 0.23479090631008148
step: 20, loss: 0.08651861548423767
step: 30, loss: 0.11929559707641602
step: 40, loss: 0.11002923548221588
step: 50, loss: 0.17901761829853058
step: 60, loss: 0.05008174479007721
step: 70, loss: 0.22240538895130157
step: 80, loss: 0.07178807258605957
step: 90, loss: 0.15054000914096832
step: 100, loss: 0.09888331592082977
step: 110, loss: 0.01598057523369789
step: 120, loss: 0.01882929727435112
step: 130, loss: 0.03407708555459976
step: 140, loss: 0.025503644719719887
step: 150, loss: 0.10935740172863007
step: 160, loss: 0.062068331986665726
step: 170, loss: 0.15642358362674713
step: 180, loss: 0.06892015784978867
step: 190, loss: 0.1699335277080536
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7192982456140352, f1=0.7422096317280454,

```

```

best_f1=0.7422096317280454
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.06078649312257767
step: 10, loss: 0.10660126060247421
step: 20, loss: 0.34233203530311584
step: 30, loss: 0.11531727015972137
step: 40, loss: 0.03834228590130806
step: 50, loss: 0.030437367036938667
step: 60, loss: 0.02929580584168434
step: 70, loss: 0.09800583869218826
step: 80, loss: 0.32523155212402344
step: 90, loss: 0.034941595047712326
step: 100, loss: 0.13637705147266388
step: 110, loss: 0.005847234278917313
step: 120, loss: 0.13901157677173615
step: 130, loss: 0.10536915808916092
step: 140, loss: 0.03124515525996685
step: 150, loss: 0.24811163544654846
step: 160, loss: 0.02259371429681778
step: 170, loss: 0.3188821077346802
step: 180, loss: 0.04741441458463669
step: 190, loss: 0.25282737612724304
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7219251336898397, f1=0.7454068241469817,
best_f1=0.7454068241469817
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05002223700284958
step: 10, loss: 0.04142037406563759
step: 20, loss: 0.04652230814099312
step: 30, loss: 0.007321993820369244
step: 40, loss: 0.052224498242139816
step: 50, loss: 0.08021646738052368
step: 60, loss: 0.0635162740945816
step: 70, loss: 0.005802797619253397
step: 80, loss: 0.014369065873324871
step: 90, loss: 0.023137178272008896
step: 100, loss: 0.009145675227046013
step: 110, loss: 0.05020295828580856

```

```

step: 120, loss: 0.01263856329023838
step: 130, loss: 0.07527519017457962
step: 140, loss: 0.029449939727783203
step: 150, loss: 0.010344588197767735
step: 160, loss: 0.0069102211855351925
step: 170, loss: 0.009993291459977627
step: 180, loss: 0.09381107985973358
step: 190, loss: 0.15801554918289185
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7309644670050761, f1=0.7525773195876289,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.02652456983923912
step: 10, loss: 0.1244487315416336
step: 20, loss: 0.019713932648301125
step: 30, loss: 0.0048160008154809475
step: 40, loss: 0.11472184211015701
step: 50, loss: 0.006632114760577679
step: 60, loss: 0.06360176205635071
step: 70, loss: 0.018893934786319733
step: 80, loss: 0.029050210490822792
step: 90, loss: 0.007375611457973719
step: 100, loss: 0.0007608954329043627
step: 110, loss: 0.14535605907440186
step: 120, loss: 0.11936677992343903
step: 130, loss: 0.042954813688993454
step: 140, loss: 0.042214635759592056
step: 150, loss: 0.01761053130030632
step: 160, loss: 0.01991255208849907
step: 170, loss: 0.10522729158401489
step: 180, loss: 0.03163396567106247
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 190, loss: 0.010433298535645008
/usr/local/lib/python3.7/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=0.7188264058679708, f1=0.7188264058679708,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```



```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004013307858258486
step: 10, loss: 0.004004590213298798
step: 20, loss: 0.0915115550160408
step: 30, loss: 0.0701693668961525
step: 40, loss: 0.006555613595992327
step: 50, loss: 0.02709578163921833
step: 60, loss: 0.004409497138112783
step: 70, loss: 0.0012708770809695125
step: 80, loss: 0.02708190120756626
step: 90, loss: 0.02570854313671589
step: 100, loss: 0.0006754864589311182
step: 110, loss: 0.013391310349106789
step: 120, loss: 0.004923022352159023
step: 130, loss: 0.1274271458387375
step: 140, loss: 0.01817917264997959
step: 150, loss: 0.05206039175391197
step: 160, loss: 0.03810686245560646
step: 170, loss: 0.006371780764311552
step: 180, loss: 0.10602252930402756
step: 190, loss: 0.008367011323571205
/usr/local/lib/python3.7/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=0.7272727272727273, f1=0.7506849315068493,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.032893478870391846
step: 10, loss: 0.004135212395340204
step: 20, loss: 0.00301779480651021
step: 30, loss: 0.07128602266311646
step: 40, loss: 0.00094379618531093
step: 50, loss: 0.003765490371733904
step: 60, loss: 0.0442904569208622
step: 70, loss: 0.024987680837512016
step: 80, loss: 0.0007194837089627981
step: 90, loss: 0.0012883262243121862
step: 100, loss: 0.0033797668293118477
step: 110, loss: 0.008999977260828018
step: 120, loss: 0.0012834318913519382
step: 130, loss: 0.0011711913393810391
step: 140, loss: 0.010537133552134037
step: 150, loss: 0.007273709401488304
step: 160, loss: 0.041374243795871735
step: 170, loss: 0.0010800486197695136
step: 180, loss: 0.006139586213976145

```

```

step: 190, loss: 0.005316128488630056
/usr/local/lib/python3.7/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=0.7164948453608248, f1=0.7586206896551724,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0008875209605321288
step: 10, loss: 0.0029812732245773077
step: 20, loss: 0.006567132193595171
step: 30, loss: 0.0010532742599025369
step: 40, loss: 0.0029186769388616085
step: 50, loss: 0.0017488928278908134
step: 60, loss: 0.0012593941064551473
step: 70, loss: 0.0003454705874901265
step: 80, loss: 0.0002724747755564749
step: 90, loss: 0.14147116243839264
step: 100, loss: 0.0022323860321193933
step: 110, loss: 0.0011916274670511484
step: 120, loss: 0.00929082278162241
step: 130, loss: 0.005604181904345751
step: 140, loss: 0.003962044138461351
step: 150, loss: 0.0026193028315901756
step: 160, loss: 0.00019176244677510113
step: 170, loss: 0.24863331019878387
step: 180, loss: 0.013925851322710514
step: 190, loss: 0.0032847432885318995
/usr/local/lib/python3.7/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=0.7197943444730078, f1=0.760705289672544,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0018202968640252948
step: 10, loss: 0.016125250607728958
step: 20, loss: 0.03736049309372902
step: 30, loss: 0.0008090147166512907
step: 40, loss: 0.013895638287067413
step: 50, loss: 0.049289681017398834
step: 60, loss: 0.01821596547961235
step: 70, loss: 0.0034383805468678474
step: 80, loss: 0.005032310727983713
step: 90, loss: 0.13039320707321167
step: 100, loss: 0.04087036848068237

```

```

step: 110, loss: 0.00824898760765791
step: 120, loss: 0.023278020322322845
step: 130, loss: 0.06823936104774475
step: 140, loss: 0.009879096411168575
step: 150, loss: 0.006940861232578754
step: 160, loss: 0.0014305270742624998
step: 170, loss: 0.0014280332252383232
step: 180, loss: 0.013269484043121338
step: 190, loss: 0.0016775712138041854
/usr/local/lib/python3.7/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=0.7297297297297297, f1=0.7750677506775067,
best_f1=0.7525773195876289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0011303828796371818
step: 10, loss: 0.002975541865453124
step: 20, loss: 0.00646353280171752
step: 30, loss: 0.002516915090382099
step: 40, loss: 0.003225281136110425
step: 50, loss: 0.0010289564961567521
step: 60, loss: 0.0008987553883343935
step: 70, loss: 0.00048029227764345706
step: 80, loss: 0.000474491564091295
step: 90, loss: 0.0018364833667874336
step: 100, loss: 0.0042603760957717896
step: 110, loss: 0.0011585673782974482
step: 120, loss: 0.0004426841624081135
step: 130, loss: 0.02793942391872406
step: 140, loss: 0.0008371319272555411
step: 150, loss: 0.0005295657319948077
step: 160, loss: 0.005217129830271006
step: 170, loss: 0.0005173876998014748
step: 180, loss: 0.010893654078245163
step: 190, loss: 0.0005830498994328082
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7455919395465995, f1=0.7712082262210797,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.

```

```

    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00038446293910965323
step: 10, loss: 0.005321075674146414
step: 20, loss: 0.0004817423177883029
step: 30, loss: 0.019917072728276253
step: 40, loss: 0.04446624964475632
step: 50, loss: 0.004777122754603624
step: 60, loss: 0.0002709657419472933
step: 70, loss: 0.0004416780429892242
step: 80, loss: 0.0009198879706673324
step: 90, loss: 0.0008404715335927904
step: 100, loss: 0.0008363716769963503
step: 110, loss: 0.0006229060236364603
step: 120, loss: 0.0003042138123419136
step: 130, loss: 0.0010626381263136864
step: 140, loss: 0.14669805765151978
step: 150, loss: 0.003723922185599804
step: 160, loss: 0.0006832635845057666
step: 170, loss: 0.001122854184359312
step: 180, loss: 0.0003621895448304713
step: 190, loss: 0.019213151186704636
/usr/local/lib/python3.7/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=0.7309644670050761, f1=0.770408163265306,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.001198428450152278
step: 10, loss: 0.018642915412783623
step: 20, loss: 0.0015223586233332753
step: 30, loss: 0.005174004007130861
step: 40, loss: 0.0004530933510977775
step: 50, loss: 0.02105294167995453
step: 60, loss: 0.0007948490092530847
step: 70, loss: 0.01676395907998085
step: 80, loss: 0.0003706621646415442
step: 90, loss: 0.0022773374803364277
step: 100, loss: 0.004760591313242912
step: 110, loss: 0.0005205973866395652
step: 120, loss: 0.012955055572092533
step: 130, loss: 0.00039067125180736184
step: 140, loss: 0.00046967979869805276
step: 150, loss: 0.00036756761255674064
step: 160, loss: 0.0015203303191810846
step: 170, loss: 0.0025559512432664633
step: 180, loss: 0.0009458505664952099

```

```

step: 190, loss: 0.012894047424197197
/usr/local/lib/python3.7/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=0.7329842931937172, f1=0.779220779220779,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004217700334265828
step: 10, loss: 0.0006394494557753205
step: 20, loss: 0.007199837360531092
step: 30, loss: 0.002015022560954094
step: 40, loss: 0.006342803128063679
step: 50, loss: 0.0014864943223074079
step: 60, loss: 0.0005027608713135123
step: 70, loss: 0.0011527311289682984
step: 80, loss: 0.0012031053192913532
step: 90, loss: 0.0009310831665061414
step: 100, loss: 0.05004461482167244
step: 110, loss: 0.0637269839644432
step: 120, loss: 0.0032173858489841223
step: 130, loss: 0.008002044633030891
step: 140, loss: 0.000668852124363184
step: 150, loss: 0.0004434005240909755
step: 160, loss: 0.002938844496384263
step: 170, loss: 0.0996476337313652
step: 180, loss: 0.0018570468528196216
step: 190, loss: 0.0003995060396846384
/usr/local/lib/python3.7/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=0.7204301075268816, f1=0.7745358090185677,
best_f1=0.7712082262210797
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006386613240465522
step: 10, loss: 0.03164894878864288
step: 20, loss: 0.0016746881883591413
step: 30, loss: 0.0005516742239706218
step: 40, loss: 0.00033620939939282835
step: 50, loss: 0.019755320623517036
step: 60, loss: 0.0005956003442406654
step: 70, loss: 0.00484848627820611
step: 80, loss: 0.01500722672790289
step: 90, loss: 0.0006225433899089694
step: 100, loss: 0.0013655164511874318

```

```

step: 110, loss: 0.001173274009488523
step: 120, loss: 0.0003721210523508489
step: 130, loss: 0.006096632685512304
step: 140, loss: 0.0022540786303579807
step: 150, loss: 0.000972093956079334
step: 160, loss: 0.0005715059814974666
step: 170, loss: 0.000813574471976608
step: 180, loss: 0.004087143111974001
step: 190, loss: 0.0005889696185477078
/usr/local/lib/python3.7/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=0.7277777777777779, f1=0.7704918032786885,
best_f1=0.7712082262210797
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, 166.77it/s]
load_f1 = 0.6812652068126521
real_f1 = 0.6526806526806527
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, 185.26it/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, 298kB/s]
Downloading: 100% 232k/232k [00:00<00:00, 2.61MB/s]
Downloading: 100% 440M/440M [00:06<00:00, 65.6MB/s]
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
```

master weights.

Defaults for this optimization level are:

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

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

After processing overrides, optimization options are:

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

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

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

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

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

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: 0.6282248497009277
step: 10, loss: 0.3820571303367615
step: 20, loss: 0.30688369274139404
step: 30, loss: 0.38028424978256226
step: 40, loss: 0.2719513773918152
step: 50, loss: 0.2762562334537506
step: 60, loss: 0.2582228183746338
step: 70, loss: 0.372784823179245
step: 80, loss: 0.3567630648612976
step: 90, loss: 0.25024470686912537
step: 100, loss: 0.2606011629104614
step: 110, loss: 0.19002613425254822
step: 120, loss: 0.24546071887016296
```



```

step: 130, loss: 0.03158487379550934
step: 140, loss: 0.1507294476032257
step: 150, loss: 0.29522082209587097
step: 160, loss: 0.11442463845014572
step: 170, loss: 0.28355932235717773
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7240506329113925, f1=0.7315914489311163,
best_f1=0.7315914489311163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.17454248666763306
step: 10, loss: 0.24831520020961761
step: 20, loss: 0.16947151720523834
step: 30, loss: 0.20299795269966125
step: 40, loss: 0.11483648419380188
step: 50, loss: 0.11814901232719421
step: 60, loss: 0.07312045246362686
step: 70, loss: 0.18052725493907928
step: 80, loss: 0.05518693849444389
step: 90, loss: 0.12992316484451294
step: 100, loss: 0.23396968841552734
step: 110, loss: 0.048873912543058395
step: 120, loss: 0.09295643121004105
step: 130, loss: 0.07698088139295578
step: 140, loss: 0.25190284848213196
step: 150, loss: 0.2572435140609741
step: 160, loss: 0.2055722326040268
step: 170, loss: 0.10633544623851776
/usr/local/lib/python3.7/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=0.7163461538461539, f1=0.7522123893805309,
best_f1=0.7315914489311163
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.2124290019273758
step: 10, loss: 0.2885635793209076
step: 20, loss: 0.040316201746463776
step: 30, loss: 0.20482584834098816
step: 40, loss: 0.12995412945747375

```

```

step: 50, loss: 0.11626067757606506
step: 60, loss: 0.08616426587104797
step: 70, loss: 0.08773665875196457
step: 80, loss: 0.05445127561688423
step: 90, loss: 0.23199620842933655
step: 100, loss: 0.06067676097154617
step: 110, loss: 0.1368797868490219
step: 120, loss: 0.06145568937063217
step: 130, loss: 0.2062867432832718
step: 140, loss: 0.16674454510211945
step: 150, loss: 0.04860633611679077
step: 160, loss: 0.07929766178131104
step: 170, loss: 0.09600520879030228
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.753623188405797, f1=0.7788018433179723,
best_f1=0.7788018433179723
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.017457270994782448
step: 10, loss: 0.242533877491951
step: 20, loss: 0.016158897429704666
step: 30, loss: 0.12285533547401428
step: 40, loss: 0.007781409192830324
step: 50, loss: 0.07659394294023514
step: 60, loss: 0.058008063584566116
step: 70, loss: 0.023321984335780144
step: 80, loss: 0.10575054585933685
step: 90, loss: 0.19329820573329926
step: 100, loss: 0.12179625034332275
step: 110, loss: 0.10604274272918701
step: 120, loss: 0.02266981080174446
step: 130, loss: 0.1538831740617752
step: 140, loss: 0.02549937553703785
step: 150, loss: 0.08062820881605148
step: 160, loss: 0.13847512006759644
step: 170, loss: 0.08746159821748734
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.7731958762886598, f1=0.7804878048780487,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.025319840759038925
step: 10, loss: 0.01712062954902649
step: 20, loss: 0.029358241707086563
step: 30, loss: 0.16358892619609833
step: 40, loss: 0.0197112038731575
step: 50, loss: 0.12407731264829636
step: 60, loss: 0.03649076074361801
step: 70, loss: 0.045820072293281555
step: 80, loss: 0.1337999701499939
step: 90, loss: 0.1934937834739685
step: 100, loss: 0.013751097023487091
step: 110, loss: 0.08402741700410843
step: 120, loss: 0.14835092425346375
step: 130, loss: 0.0713755264878273
step: 140, loss: 0.010899960063397884
step: 150, loss: 0.047431580722332
step: 160, loss: 0.24122317135334015
step: 170, loss: 0.00779977859929204
/usr/local/lib/python3.7/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=0.772972972972973, f1=0.775, best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.030111026018857956
step: 10, loss: 0.00853356160223484
step: 20, loss: 0.026149030774831772
step: 30, loss: 0.07743509113788605
step: 40, loss: 0.010131125338375568
step: 50, loss: 0.13515496253967285
step: 60, loss: 0.12510401010513306
step: 70, loss: 0.08180789649486542
step: 80, loss: 0.02499382384121418
step: 90, loss: 0.004628363065421581
step: 100, loss: 0.006642083637416363
step: 110, loss: 0.00195550755597651
step: 120, loss: 0.0161668099462986
step: 130, loss: 0.011173772625625134
step: 140, loss: 0.015179879032075405
step: 150, loss: 0.07868611812591553

```

```

step: 160, loss: 0.11577768623828888
step: 170, loss: 0.011886738240718842
/usr/local/lib/python3.7/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=0.7608695652173912, f1=0.7552083333333334,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0030600279569625854
step: 10, loss: 0.002759203314781189
step: 20, loss: 0.0026660196017473936
step: 30, loss: 0.02776496484875679
step: 40, loss: 0.004751617554575205
step: 50, loss: 0.03879593685269356
step: 60, loss: 0.006356813479214907
step: 70, loss: 0.0061910077929496765
step: 80, loss: 0.008898979984223843
step: 90, loss: 0.00041196070378646255
step: 100, loss: 0.028146175667643547
step: 110, loss: 0.03133145347237587
step: 120, loss: 0.0027599127497524023
step: 130, loss: 0.3163331151008606
step: 140, loss: 0.0028377140406519175
step: 150, loss: 0.013117795810103416
step: 160, loss: 0.003439904423430562
step: 170, loss: 0.08863722532987595
/usr/local/lib/python3.7/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=0.7447916666666667, f1=0.7611940298507462,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.010725390166044235
step: 10, loss: 0.008487322367727757
step: 20, loss: 0.001621211995370686
step: 30, loss: 0.027261173352599144
step: 40, loss: 0.0008816521149128675
step: 50, loss: 0.007194503676146269
step: 60, loss: 0.030436914414167404
step: 70, loss: 0.011250643990933895
step: 80, loss: 0.005541532766073942
step: 90, loss: 0.02672652155160904
step: 100, loss: 0.09522170573472977
step: 110, loss: 0.045716818422079086

```

```

step: 120, loss: 0.03921923786401749
step: 130, loss: 0.045025069266557693
step: 140, loss: 0.011015730910003185
step: 150, loss: 0.12641146779060364
step: 160, loss: 0.054140664637088776
step: 170, loss: 0.003097235457971692
/usr/local/lib/python3.7/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=0.7422680412371134, f1=0.7609756097560975,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.003318576840683818
step: 10, loss: 0.010014811530709267
step: 20, loss: 0.0015953043475747108
step: 30, loss: 0.07498791813850403
step: 40, loss: 0.08021169155836105
step: 50, loss: 0.002198456786572933
step: 60, loss: 0.007459302432835102
step: 70, loss: 0.06645450741052628
step: 80, loss: 0.00947557482868433
step: 90, loss: 0.07866921275854111
step: 100, loss: 0.013034412637352943
step: 110, loss: 0.0027232717256993055
step: 120, loss: 0.033557746559381485
step: 130, loss: 0.11538013070821762
step: 140, loss: 0.004078416153788567
step: 150, loss: 0.017846936360001564
step: 160, loss: 0.02682776376605034
step: 170, loss: 0.06308139860630035
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=0.7611940298507462, f1=0.7692307692307692,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.11821132153272629
step: 10, loss: 0.001950261415913701
step: 20, loss: 0.024629101157188416
step: 30, loss: 0.008167391642928123
step: 40, loss: 0.10482414066791534
step: 50, loss: 0.08426757156848907
step: 60, loss: 0.00446031428873539

```

```

step: 70, loss: 0.010079209692776203
step: 80, loss: 0.00243981066159904
step: 90, loss: 0.0033648137468844652
step: 100, loss: 0.0004062523366883397
step: 110, loss: 0.021684372797608376
step: 120, loss: 0.0014780573546886444
step: 130, loss: 0.0018274747999385
step: 140, loss: 0.026687216013669968
step: 150, loss: 0.010769705288112164
step: 160, loss: 0.0013454610016196966
step: 170, loss: 0.000859213003423065
/usr/local/lib/python3.7/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=0.75, f1=0.7673860911270983, best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.05263297259807587
step: 10, loss: 0.001451238989830017
step: 20, loss: 0.04501296579837799
step: 30, loss: 0.00026734848506748676
step: 40, loss: 0.0024263744708150625
step: 50, loss: 0.0024828764144331217
step: 60, loss: 0.031704291701316833
step: 70, loss: 0.004000208806246519
step: 80, loss: 0.0006643740925937891
step: 90, loss: 0.0009838631376624107
step: 100, loss: 0.005930631421506405
step: 110, loss: 0.03520653396844864
step: 120, loss: 0.027372922748327255
step: 130, loss: 0.000644715444650501
step: 140, loss: 0.0003416293766349554
step: 150, loss: 0.04957769066095352
step: 160, loss: 0.008984719403088093
step: 170, loss: 0.007320203818380833
/usr/local/lib/python3.7/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=0.7453083109919573, f1=0.7700258397932818,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009798840619623661
step: 10, loss: 0.001533341477625072
step: 20, loss: 0.00036513348459266126
step: 30, loss: 0.005877891555428505

```

```

step: 40, loss: 0.003430369310081005
step: 50, loss: 0.0009122680057771504
step: 60, loss: 0.0653616338968277
step: 70, loss: 0.12522801756858826
step: 80, loss: 0.0004261146823409945
step: 90, loss: 0.00917025189846754
step: 100, loss: 0.001363275689072907
step: 110, loss: 0.006026356481015682
step: 120, loss: 0.06634782999753952
step: 130, loss: 0.06155085191130638
step: 140, loss: 0.015953801572322845
step: 150, loss: 0.009153414517641068
step: 160, loss: 0.006962120532989502
step: 170, loss: 0.0006493252003565431
/usr/local/lib/python3.7/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=0.7616580310880828, f1=0.7808564231738037,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004735400900244713
step: 10, loss: 0.0026789596304297447
step: 20, loss: 0.0038326061330735683
step: 30, loss: 0.0009733536280691624
step: 40, loss: 0.14263497292995453
step: 50, loss: 0.004585574381053448
step: 60, loss: 0.0055251410230994225
step: 70, loss: 0.0016049528494477272
step: 80, loss: 0.08292330056428909
step: 90, loss: 0.0008202476892620325
step: 100, loss: 0.04888594523072243
step: 110, loss: 0.006725529674440622
step: 120, loss: 0.017615996301174164
step: 130, loss: 0.0007776060956530273
step: 140, loss: 0.0016999561339616776
step: 150, loss: 0.0013985224068164825
step: 160, loss: 0.0036377795040607452
step: 170, loss: 0.012030177749693394
/usr/local/lib/python3.7/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=0.7538461538461538, f1=0.7669172932330827,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")

```

```

step: 0, loss: 0.004305372480303049
step: 10, loss: 0.0006036568083800375
step: 20, loss: 0.0005120151326991618
step: 30, loss: 0.0021379271056503057
step: 40, loss: 0.002469244645908475
step: 50, loss: 0.24916177988052368
step: 60, loss: 0.01537253800779581
step: 70, loss: 0.006324264220893383
step: 80, loss: 0.0016414436977356672
step: 90, loss: 0.0006632033619098365
step: 100, loss: 0.018117636442184448
step: 110, loss: 0.002440334064885974
step: 120, loss: 0.00960470549762249
step: 130, loss: 0.0006607882678508759
step: 140, loss: 0.0009716842323541641
step: 150, loss: 0.0005770110874436796
step: 160, loss: 0.004161078017205
step: 170, loss: 0.0015841309214010835
/usr/local/lib/python3.7/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=0.7700258397932818,
best_f1=0.7804878048780487
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0002805921249091625
step: 10, loss: 0.003237007884308696
step: 20, loss: 0.1292945295572281
step: 30, loss: 0.0004776118730660528
step: 40, loss: 0.005232604686170816
step: 50, loss: 0.007661826908588409
step: 60, loss: 0.11941757053136826
step: 70, loss: 0.0007561847451142967
step: 80, loss: 0.003127575619146228
step: 90, loss: 0.0006350639159791172
step: 100, loss: 0.0018022560980170965
step: 110, loss: 0.0003326215664856136
step: 120, loss: 0.002977433381602168
step: 130, loss: 0.001593229710124433
step: 140, loss: 0.0062127625569701195
step: 150, loss: 0.00033644732320681214
step: 160, loss: 0.050782959908246994
step: 170, loss: 0.004557653795927763
/usr/local/lib/python3.7/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=0.7613941018766757, f1=0.7712082262210797,

```



```

best_f1=0.7804878048780487
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level O2:  FP16 training with FP32 batchnorm and FP32
master weights.

Defaults for this optimization level are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : O2
cast_model_type        : torch.float16
patch_torch_functions  : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning:  multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
Oit [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
1916it [00:06, 275.85it/s]
load_f1 = 0.4511627906976744
real_f1 = 0.448512585812357
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, 248.75it/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: 0.6059665679931641
step: 10, loss: 0.6404613852500916
step: 20, loss: 0.4676644206047058
step: 30, loss: 0.26009389758110046
step: 40, loss: 0.1380099654197693
step: 50, loss: 0.07081162929534912
step: 60, loss: 0.12009942531585693
step: 70, loss: 0.14359523355960846
step: 80, loss: 0.08826525509357452
step: 90, loss: 0.15043465793132782
step: 100, loss: 0.003984111826866865
step: 110, loss: 0.2046249806880951
step: 120, loss: 0.021998027339577675
step: 130, loss: 0.26383280754089355
step: 140, loss: 0.016354447230696678
step: 150, loss: 0.11166511476039886
step: 160, loss: 0.00678722420707345
step: 170, loss: 0.04700412228703499
step: 180, loss: 0.018539119511842728
step: 190, loss: 0.01659712754189968
step: 200, loss: 0.08994735777378082
step: 210, loss: 0.016111386939883232
step: 220, loss: 0.024780653417110443
step: 230, loss: 0.11063911020755768
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:

```

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=0.9576837416481068, f1=0.962962962962963,

best_f1=0.962962962962963

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

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

step: 0, loss: 0.01464009564369917

step: 10, loss: 0.007374242413789034

step: 20, loss: 0.2673022449016571

step: 30, loss: 0.11102664470672607

step: 40, loss: 0.10980721563100815

step: 50, loss: 0.008010907098650932

step: 60, loss: 0.0030019194819033146

step: 70, loss: 0.0338609479367733

step: 80, loss: 0.0016365477349609137

step: 90, loss: 0.11316360533237457

step: 100, loss: 0.00510804820805788

step: 110, loss: 0.17442892491817474

step: 120, loss: 0.13757480680942535

step: 130, loss: 0.029067570343613625

step: 140, loss: 0.004009446129202843

step: 150, loss: 0.010338693857192993

step: 160, loss: 0.02354883775115013

step: 170, loss: 0.028716951608657837

step: 180, loss: 0.024133404716849327

step: 190, loss: 0.005715614650398493

step: 200, loss: 0.002703494392335415

step: 210, loss: 0.007546085398644209

step: 220, loss: 0.17336145043373108

step: 230, loss: 0.03036997653543949

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

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

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

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=0.9700996677740864, f1=0.9634551495016611,

best_f1=0.9634551495016611

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

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

step: 0, loss: 0.01562969572842121

step: 10, loss: 0.010025249794125557

step: 20, loss: 0.034901730716228485

step: 30, loss: 0.002894091885536909

```

step: 40, loss: 0.02577958069741726
step: 50, loss: 0.02769869565963745
step: 60, loss: 0.03202149271965027
step: 70, loss: 0.0036791888996958733
step: 80, loss: 0.0029583412688225508
step: 90, loss: 0.11965889483690262
step: 100, loss: 0.0023448735009878874
step: 110, loss: 0.0010150948073714972
step: 120, loss: 0.008106199093163013
step: 130, loss: 0.0010735881514847279
step: 140, loss: 0.016672976315021515
step: 150, loss: 0.09489310532808304
step: 160, loss: 0.14365561306476593
step: 170, loss: 0.026025768369436264
step: 180, loss: 0.03778831660747528
step: 190, loss: 0.024897294119000435
step: 200, loss: 0.04490743204951286
step: 210, loss: 0.08929906040430069
step: 220, loss: 0.05385292321443558
step: 230, loss: 0.000660375808365643
/usr/local/lib/python3.7/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=0.9695603156708005, f1=0.96045197740113,
best_f1=0.9634551495016611
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.002160192234441638
step: 10, loss: 0.005252293311059475
step: 20, loss: 0.0014124322915449739
step: 30, loss: 0.0008963234140537679
step: 40, loss: 0.030405765399336815
step: 50, loss: 0.00039758719503879547
step: 60, loss: 0.002599171129986644
step: 70, loss: 0.015827715396881104
step: 80, loss: 0.004400173202157021
step: 90, loss: 0.012289407663047314
step: 100, loss: 0.0019511538557708263
step: 110, loss: 0.0070632100105285645
step: 120, loss: 0.02369796670973301
step: 130, loss: 0.009322659112513065
step: 140, loss: 0.07659043371677399
step: 150, loss: 0.018541615456342697
step: 160, loss: 0.006357137113809586
step: 170, loss: 0.007736450992524624
step: 180, loss: 0.0009204774396494031
step: 190, loss: 0.013496620580554008

```

```

step: 200, loss: 0.05411477014422417
step: 210, loss: 0.08828277885913849
step: 220, loss: 0.0055483500473201275
step: 230, loss: 0.012469667010009289
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9720670391061451, f1=0.9675977653631285,
best_f1=0.9675977653631285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.009438898414373398
step: 10, loss: 0.003041165880858898
step: 20, loss: 0.01492415089160204
step: 30, loss: 0.0003634895256254822
step: 40, loss: 0.0007453294238075614
step: 50, loss: 0.0005653211264871061
step: 60, loss: 0.15008406341075897
step: 70, loss: 0.008538118563592434
step: 80, loss: 0.002334153512492776
step: 90, loss: 0.0013679706025868654
step: 100, loss: 0.0005269358516670763
step: 110, loss: 0.0003310497268103063
step: 120, loss: 0.0002329212293261662
step: 130, loss: 0.03271283954381943
step: 140, loss: 0.01712079532444477
step: 150, loss: 0.0011333232978358865
step: 160, loss: 0.000377634190954268
step: 170, loss: 0.046493012458086014
step: 180, loss: 0.0009630703716538846
step: 190, loss: 0.12122143059968948
step: 200, loss: 0.07485862076282501
step: 210, loss: 0.01831107586622238
step: 220, loss: 0.060126811265945435
step: 230, loss: 0.011954915709793568
/usr/local/lib/python3.7/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=0.9665924276169264, f1=0.953125, best_f1=0.9675977653631285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.033643919974565506

```

```

step: 10, loss: 0.0015716498019173741
step: 20, loss: 0.004672353155910969
step: 30, loss: 0.004044719971716404
step: 40, loss: 0.0011746634263545275
step: 50, loss: 0.02020263671875
step: 60, loss: 0.06234704330563545
step: 70, loss: 0.015422842465341091
step: 80, loss: 0.005512411240488291
step: 90, loss: 0.016706688329577446
step: 100, loss: 0.03583085909485817
step: 110, loss: 0.036874864250421524
step: 120, loss: 0.009359107352793217
step: 130, loss: 0.0004215784720145166
step: 140, loss: 0.0010565277189016342
step: 150, loss: 0.004057268146425486
step: 160, loss: 0.0035414351150393486
step: 170, loss: 0.00030566996429115534
step: 180, loss: 0.03699813783168793
step: 190, loss: 0.0012059350265190005
step: 200, loss: 0.005312207154929638
step: 210, loss: 0.00014439577353186905
step: 220, loss: 0.00019234228238929063
step: 230, loss: 0.029930923134088516
/usr/local/lib/python3.7/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=0.9678135405105438, f1=0.9664429530201343,
best_f1=0.9675977653631285
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0023730508983135223
step: 10, loss: 0.0026272027753293514
step: 20, loss: 0.0018624268705025315
step: 30, loss: 0.030708065256476402
step: 40, loss: 0.00489537138491869
step: 50, loss: 0.00042081845458596945
step: 60, loss: 0.002824316034093499
step: 70, loss: 0.0017781464848667383
step: 80, loss: 0.001766740926541388
step: 90, loss: 0.00040264346171170473
step: 100, loss: 0.001479164231568575
step: 110, loss: 0.0013787953648716211
step: 120, loss: 0.0010553153697401285
step: 130, loss: 0.06271138787269592
step: 140, loss: 0.0001714264799375087
step: 150, loss: 0.003046442288905382
step: 160, loss: 0.06339448690414429

```

```

step: 170, loss: 0.0033410692121833563
step: 180, loss: 0.004350162576884031
step: 190, loss: 0.0005433941842056811
step: 200, loss: 0.03151823952794075
step: 210, loss: 0.00013163995754439384
step: 220, loss: 0.03537433221936226
step: 230, loss: 0.0015197627944871783
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9744160177975528, f1=0.9594594594594594,
best_f1=0.9594594594594594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004009914118796587
step: 10, loss: 0.00018978433217853308
step: 20, loss: 0.00747944600880146
step: 30, loss: 0.0009242534288205206
step: 40, loss: 0.06910267472267151
step: 50, loss: 0.0005853584152646363
step: 60, loss: 0.0001173363853013143
step: 70, loss: 0.0010681254789233208
step: 80, loss: 0.0009250364964827895
step: 90, loss: 0.0001890470739454031
step: 100, loss: 0.008949533104896545
step: 110, loss: 0.0033564630430191755
step: 120, loss: 0.04808538034558296
step: 130, loss: 0.001680124318227172
step: 140, loss: 0.00021497583657037467
step: 150, loss: 0.00019692082423716784
step: 160, loss: 0.00010224981815554202
step: 170, loss: 0.007910053245723248
step: 180, loss: 0.0006775293732061982
step: 190, loss: 0.0015575801953673363
step: 200, loss: 0.0012878201669082046
step: 210, loss: 0.002923607360571623
step: 220, loss: 0.0003843313897959888
step: 230, loss: 0.00034746824530884624
/usr/local/lib/python3.7/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=0.972972972972973, f1=0.9627118644067796,
best_f1=0.9594594594594594

```



```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.000133520967210643
step: 10, loss: 0.00031968214898370206
step: 20, loss: 0.005085027310997248
step: 30, loss: 8.979460108093917e-05
step: 40, loss: 0.004925978370010853
step: 50, loss: 9.756039071362466e-05
step: 60, loss: 0.00011222836474189535
step: 70, loss: 0.0026877750642597675
step: 80, loss: 0.0003419673885218799
step: 90, loss: 0.0003066405770368874
step: 100, loss: 0.0005112490034662187
step: 110, loss: 7.679578266106546e-05
step: 120, loss: 6.78139113006182e-05
step: 130, loss: 0.00013889594993088394
step: 140, loss: 5.4140771680977196e-05
step: 150, loss: 0.0007886136882007122
step: 160, loss: 7.312455272767693e-05
step: 170, loss: 6.867134652566165e-05
step: 180, loss: 0.0031838323920965195
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 190, loss: 0.00017814284365158528
step: 200, loss: 0.000639508361928165
step: 210, loss: 0.0020414958707988262
step: 220, loss: 9.686785051599145e-05
step: 230, loss: 0.0001606520381756127
/usr/local/lib/python3.7/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=0.9704545454545453, f1=0.954233409610984,
best_f1=0.9594594594594594
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00013268926704768091
step: 10, loss: 8.840108057484031e-05
step: 20, loss: 7.30559549992904e-05
step: 30, loss: 6.0211717936908826e-05
step: 40, loss: 0.005582879763096571
step: 50, loss: 0.013344711624085903
step: 60, loss: 0.0007766852504573762
step: 70, loss: 0.001950457226485014
step: 80, loss: 5.0787108193617314e-05
step: 90, loss: 0.0001129716620198451
step: 100, loss: 7.048757834127173e-05
step: 110, loss: 0.006471714470535517

```

```

step: 120, loss: 5.371864972403273e-05
step: 130, loss: 0.0005583566380664706
step: 140, loss: 0.0477508120238781
step: 150, loss: 0.027626564726233482
step: 160, loss: 0.00013078092888463289
step: 170, loss: 4.155694477958605e-05
step: 180, loss: 0.00464747054502368
step: 190, loss: 0.005307209677994251
step: 200, loss: 3.385449235793203e-05
step: 210, loss: 0.0013157343491911888
step: 220, loss: 4.961088052368723e-05
step: 230, loss: 0.00028239141101948917
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9764837625979844, f1=0.9606299212598425,
best_f1=0.9606299212598425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004844742361456156
step: 10, loss: 0.0001295085676247254
step: 20, loss: 0.0009309934102930129
step: 30, loss: 0.058469660580158234
step: 40, loss: 6.216600013431162e-05
step: 50, loss: 0.00011094992805738002
step: 60, loss: 0.0006607670220546424
step: 70, loss: 7.172777986852452e-05
step: 80, loss: 3.9805912820156664e-05
step: 90, loss: 0.04817216470837593
step: 100, loss: 0.00023055882775224745
step: 110, loss: 0.0011924298014491796
step: 120, loss: 0.0025864322669804096
step: 130, loss: 8.969767077360302e-05
step: 140, loss: 0.000369323039194569
step: 150, loss: 0.016088994219899178
step: 160, loss: 0.0002556624822318554
step: 170, loss: 0.016479896381497383
step: 180, loss: 0.004941195715218782
step: 190, loss: 0.00012007569603156298
step: 200, loss: 0.0022529461421072483
step: 210, loss: 0.00015553658886346966
step: 220, loss: 5.567157131736167e-05
step: 230, loss: 5.821090599056333e-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=0.9741863075196409, f1=0.9638009049773756,
best_f1=0.9606299212598425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 8.953742508310825e-05
step: 10, loss: 0.00012363515270408243
step: 20, loss: 0.0048257820308208466
step: 30, loss: 0.0003420225693844259
step: 40, loss: 0.000914964999537915
step: 50, loss: 0.0034008612856268883
step: 60, loss: 0.0001946217380464077
step: 70, loss: 0.0039076171815395355
step: 80, loss: 0.0001765187334967777
step: 90, loss: 3.92262518289499e-05
step: 100, loss: 5.565587343880907e-05
step: 110, loss: 0.00011285937944194302
step: 120, loss: 5.867262007086538e-05
step: 130, loss: 0.0001426342932973057
step: 140, loss: 0.00016118933854158968
step: 150, loss: 0.00023972959024831653
step: 160, loss: 0.00018672940495889634
step: 170, loss: 0.00023758895986247808
step: 180, loss: 0.0015851451316848397
step: 190, loss: 0.00010999779624398798
step: 200, loss: 8.693575364304706e-05
step: 210, loss: 7.476213068002835e-05
step: 220, loss: 0.0035878769122064114
step: 230, loss: 0.0066480315290391445
/usr/local/lib/python3.7/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=0.968609865470852, f1=0.9581920903954803,
best_f1=0.9606299212598425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0001435736776329577
step: 10, loss: 0.001577515620738268
step: 20, loss: 8.064211579039693e-05
step: 30, loss: 0.00022122752852737904
step: 40, loss: 9.904876060318202e-05
step: 50, loss: 0.000190642531379126
step: 60, loss: 0.002532264683395624
step: 70, loss: 0.0003399299457669258

```

```

step: 80, loss: 0.004071770701557398
step: 90, loss: 0.000709510175511241
step: 100, loss: 8.803256059763953e-05
step: 110, loss: 0.002418093616142869
step: 120, loss: 0.02491261251270771
step: 130, loss: 0.00018116930732503533
step: 140, loss: 0.0002108907065121457
step: 150, loss: 5.563598097069189e-05
step: 160, loss: 0.00017684789781924337
step: 170, loss: 0.007233056239783764
step: 180, loss: 6.16183242527768e-05
step: 190, loss: 8.479761891067028e-05
step: 200, loss: 0.0008639158331789076
step: 210, loss: 3.745682624867186e-05
step: 220, loss: 0.0002802997187245637
step: 230, loss: 0.0070289866998791695
/usr/local/lib/python3.7/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=0.9753914988814317, f1=0.963882618510158,
best_f1=0.9606299212598425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00011892592010553926
step: 10, loss: 0.0038785110227763653
step: 20, loss: 0.0031633591279387474
step: 30, loss: 0.001306579913944006
step: 40, loss: 0.00015858470578677952
step: 50, loss: 0.0002133284433512017
step: 60, loss: 7.776783604640514e-05
step: 70, loss: 0.00027437068638391793
step: 80, loss: 6.646446126978844e-05
step: 90, loss: 0.0001817553857108578
step: 100, loss: 9.511989628663287e-05
step: 110, loss: 0.10258117318153381
step: 120, loss: 3.43090869137086e-05
step: 130, loss: 8.786627586232498e-05
step: 140, loss: 9.552011761115864e-05
step: 150, loss: 5.917467569815926e-05
step: 160, loss: 0.0021105040796101093
step: 170, loss: 6.198737537488341e-05
step: 180, loss: 0.00012047730706399307
step: 190, loss: 6.495473644463345e-05
step: 200, loss: 3.517682489473373e-05
step: 210, loss: 0.0018499696161597967
step: 220, loss: 0.00041394151048734784
step: 230, loss: 0.003433352569118142

```

```

/usr/local/lib/python3.7/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=0.9752252252252253, f1=0.9603624009060023,
best_f1=0.9606299212598425
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 7.477695908164605e-05
step: 10, loss: 7.884634396759793e-05
step: 20, loss: 6.377466343110427e-05
step: 30, loss: 0.00015348794113378972
step: 40, loss: 0.00012688404240179807
step: 50, loss: 0.0326264388859272
step: 60, loss: 0.0001559123193146661
step: 70, loss: 0.0008347387774847448
step: 80, loss: 0.01020828913897276
step: 90, loss: 0.0003663110837806016
step: 100, loss: 7.735345570836216e-05
step: 110, loss: 3.610464045777917e-05
step: 120, loss: 0.0003101910697296262
step: 130, loss: 6.442405720008537e-05
step: 140, loss: 6.720339297316968e-05
step: 150, loss: 0.00015907138003967702
step: 160, loss: 5.80556588829495e-05
step: 170, loss: 4.347569483797997e-05
step: 180, loss: 0.0033641294576227665
step: 190, loss: 0.0011643711477518082
step: 200, loss: 0.00017331539129372686
step: 210, loss: 9.8819138656836e-05
step: 220, loss: 0.0006090162205509841
step: 230, loss: 9.00161248864606e-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=0.9741863075196409, f1=0.96045197740113,
best_f1=0.9606299212598425
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

```

Defaults for this optimization level are:

```

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

```

```

loss_scale          : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled             : True
opt_level           : 02
cast_model_type     : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32  : True
master_weights      : True
loss_scale          : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
2473it [00:12, 204.99it/s]
load_f1 = 0.9732739420935412
real_f1 = 0.9732142857142857
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, 249.46it/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: 0.61080002784729
step: 10, loss: 0.5383345484733582
step: 20, loss: 0.5422026515007019
step: 30, loss: 0.24897408485412598
step: 40, loss: 0.30064016580581665
step: 50, loss: 0.15027444064617157
step: 60, loss: 0.11627935618162155
step: 70, loss: 0.11498306691646576
step: 80, loss: 0.07837162166833878
step: 90, loss: 0.35957029461860657
step: 100, loss: 0.07693329453468323
step: 110, loss: 0.15692129731178284
step: 120, loss: 0.18612031638622284
step: 130, loss: 0.11197390407323837
step: 140, loss: 0.2053571343421936
step: 150, loss: 0.09475703537464142
step: 160, loss: 0.023135747760534286
step: 170, loss: 0.27847379446029663
step: 180, loss: 0.1729660928249359
step: 190, loss: 0.013136213645339012
step: 200, loss: 0.1757017970085144
step: 210, loss: 0.07172787934541702
step: 220, loss: 0.20855997502803802
step: 230, loss: 0.13573120534420013
step: 240, loss: 0.10399205982685089
step: 250, loss: 0.08755986392498016
step: 260, loss: 0.1352640688419342
step: 270, loss: 0.02718198671936989
step: 280, loss: 0.0708976536989212
step: 290, loss: 0.2047608494758606
step: 300, loss: 0.06678146123886108
step: 310, loss: 0.2699015438556671
step: 320, loss: 0.12573407590389252
step: 330, loss: 0.12598513066768646
step: 340, loss: 0.13168513774871826
step: 350, loss: 0.15232130885124207
step: 360, loss: 0.09870448708534241
step: 370, loss: 0.15023164451122284
step: 380, loss: 0.01960039511322975
step: 390, loss: 0.17255382239818573
step: 400, loss: 0.25480735301971436
step: 410, loss: 0.043844617903232574
```



```

step: 420, loss: 0.08254163712263107
step: 430, loss: 0.16303107142448425
step: 440, loss: 0.03942703828215599
step: 450, loss: 0.019924890249967575
step: 460, loss: 0.029005005955696106
step: 470, loss: 0.12382644414901733
step: 480, loss: 0.036348674446344376
step: 490, loss: 0.1035243347287178
step: 500, loss: 0.08613231033086777
step: 510, loss: 0.08271219581365585
step: 520, loss: 0.05813751742243767
step: 530, loss: 0.009291013702750206
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9163972286374134, f1=0.917516218721038,
best_f1=0.917516218721038
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.20204319059848785
step: 10, loss: 0.1262097954750061
step: 20, loss: 0.0921221598982811
step: 30, loss: 0.15858756005764008
step: 40, loss: 0.10254906117916107
step: 50, loss: 0.19348230957984924
step: 60, loss: 0.051150087267160416
step: 70, loss: 0.09015528112649918
step: 80, loss: 0.1060400903224945
step: 90, loss: 0.021531781181693077
step: 100, loss: 0.0878034383058548
step: 110, loss: 0.02726954221725464
step: 120, loss: 0.24273303151130676
step: 130, loss: 0.08173400908708572
step: 140, loss: 0.051366906613111496
step: 150, loss: 0.0993482768535614
step: 160, loss: 0.1513783037662506
step: 170, loss: 0.06704898923635483
step: 180, loss: 0.054891109466552734
step: 190, loss: 0.07453974336385727
step: 200, loss: 0.03488297760486603
step: 210, loss: 0.05707955360412598
step: 220, loss: 0.07237949967384338
step: 230, loss: 0.038091715425252914

```

```

step: 240, loss: 0.06202524155378342
step: 250, loss: 0.01808134652674198
step: 260, loss: 0.008136662654578686
step: 270, loss: 0.27872347831726074
step: 280, loss: 0.07715898007154465
step: 290, loss: 0.07250301539897919
step: 300, loss: 0.1378878355026245
step: 310, loss: 0.013047021813690662
step: 320, loss: 0.2227075695991516
step: 330, loss: 0.10023956745862961
step: 340, loss: 0.04766424745321274
step: 350, loss: 0.0017953314818441868
step: 360, loss: 0.04951445385813713
step: 370, loss: 0.1576216220855713
step: 380, loss: 0.04148713871836662
step: 390, loss: 0.14564065635204315
step: 400, loss: 0.014297065325081348
step: 410, loss: 0.04907683655619621
step: 420, loss: 0.09385883063077927
step: 430, loss: 0.020085565745830536
step: 440, loss: 0.0867990255355835
step: 450, loss: 0.06615625321865082
step: 460, loss: 0.018882635980844498
step: 470, loss: 0.013921651057898998
step: 480, loss: 0.19423410296440125
step: 490, loss: 0.04777717962861061
step: 500, loss: 0.26497894525527954
step: 510, loss: 0.03894959017634392
step: 520, loss: 0.1290690153837204
step: 530, loss: 0.08920522779226303
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9197012138188609, f1=0.9212233549582948,
best_f1=0.9212233549582948
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.09963775426149368
step: 10, loss: 0.06658117473125458
step: 20, loss: 0.06964298337697983
step: 30, loss: 0.04161316901445389
step: 40, loss: 0.015757402405142784
step: 50, loss: 0.16876521706581116

```

step: 60, loss: 0.04282680153846741
step: 70, loss: 0.0025942493230104446
step: 80, loss: 0.002261054702103138
step: 90, loss: 0.008065884932875633
step: 100, loss: 0.019018882885575294
step: 110, loss: 0.0031959409825503826
step: 120, loss: 0.008375609293580055
step: 130, loss: 0.008090573363006115
step: 140, loss: 0.007735856808722019
step: 150, loss: 0.018580669537186623
step: 160, loss: 0.026295211166143417
step: 170, loss: 0.13152463734149933
step: 180, loss: 0.0281673576682806
step: 190, loss: 0.012579384259879589
step: 200, loss: 0.01760731264948845
step: 210, loss: 0.031072836369276047
step: 220, loss: 0.0322861485183239
step: 230, loss: 0.09881209582090378
step: 240, loss: 0.017911860719323158
step: 250, loss: 0.05792003870010376
step: 260, loss: 0.028834166005253792
step: 270, loss: 0.006576668005436659
step: 280, loss: 0.04272131621837616
step: 290, loss: 0.006678635720163584
step: 300, loss: 0.006574713625013828
step: 310, loss: 0.09028933197259903
step: 320, loss: 0.01748690754175186
step: 330, loss: 0.02259197272360325
step: 340, loss: 0.007899956777691841
step: 350, loss: 0.20731481909751892
step: 360, loss: 0.06606235355138779
step: 370, loss: 0.03412339463829994
step: 380, loss: 0.009790303185582161
step: 390, loss: 0.02077101729810238
step: 400, loss: 0.03901120275259018
step: 410, loss: 0.018152492120862007
step: 420, loss: 0.29863956570625305
step: 430, loss: 0.18568851053714752
step: 440, loss: 0.02233223058283329
step: 450, loss: 0.09131369739770889
step: 460, loss: 0.07400676608085632
step: 470, loss: 0.13872259855270386
step: 480, loss: 0.006677550729364157
step: 490, loss: 0.006065468303859234
step: 500, loss: 0.1174003928899765
step: 510, loss: 0.010079119354486465
step: 520, loss: 0.10296302288770676
step: 530, loss: 0.2978144586086273

```

/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9231468849477036, f1=0.9223785746709033,
best_f1=0.9223785746709033
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.013866585679352283
step: 10, loss: 0.0033226991072297096
step: 20, loss: 0.005799619946628809
step: 30, loss: 0.01471307035535574
step: 40, loss: 0.042139846831560135
step: 50, loss: 0.045392412692308426
step: 60, loss: 0.019685864448547363
step: 70, loss: 0.010254059918224812
step: 80, loss: 0.012907734140753746
step: 90, loss: 0.14884167909622192
step: 100, loss: 0.008325847797095776
step: 110, loss: 0.16734108328819275
step: 120, loss: 0.003086691489443183
step: 130, loss: 0.00670077558606863
step: 140, loss: 0.00750345503911376
step: 150, loss: 0.07432851940393448
step: 160, loss: 0.07939630001783371
step: 170, loss: 0.0044620041735470295
step: 180, loss: 0.10577023029327393
step: 190, loss: 0.04224696755409241
step: 200, loss: 0.046906013041734695
step: 210, loss: 0.09913991391658783
step: 220, loss: 0.04878959804773331
step: 230, loss: 0.190564826130867
step: 240, loss: 0.016304923221468925
step: 250, loss: 0.04683208838105202
step: 260, loss: 0.010102048516273499
step: 270, loss: 0.011525842361152172
step: 280, loss: 0.03712131455540657
step: 290, loss: 0.02570868283510208
step: 300, loss: 0.0012380675179883838
step: 310, loss: 0.0034196788910776377
step: 320, loss: 0.014552201144397259
step: 330, loss: 0.0279754176735878
step: 340, loss: 0.1538609117269516
step: 350, loss: 0.08477845788002014

```

```

step: 360, loss: 0.02728867344558239
step: 370, loss: 0.006539309863001108
step: 380, loss: 0.00952373631298542
step: 390, loss: 0.025081831961870193
step: 400, loss: 0.007101984228938818
step: 410, loss: 0.0015666221734136343
step: 420, loss: 0.020426135510206223
step: 430, loss: 0.10110122710466385
step: 440, loss: 0.017600443214178085
step: 450, loss: 0.008612890727818012
step: 460, loss: 0.11631982028484344
step: 470, loss: 0.08710470050573349
step: 480, loss: 0.13474102318286896
step: 490, loss: 0.026089996099472046
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 500, loss: 0.005323675461113453
step: 510, loss: 0.04967121034860611
step: 520, loss: 0.06416371464729309
step: 530, loss: 0.05516371130943298
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
/usr/local/lib/python3.7/dist-packages/torch/optim/lr_scheduler.py:216:
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=0.9232209737827715, f1=0.9215777262180975,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.19147244095802307
step: 10, loss: 0.04553268477320671
step: 20, loss: 0.02001550793647766
step: 30, loss: 0.0027934922836720943
step: 40, loss: 0.348354697227478
step: 50, loss: 0.003283468075096607
step: 60, loss: 0.09874971956014633
step: 70, loss: 0.00728023424744606
step: 80, loss: 0.0057103391736745834
step: 90, loss: 0.0012580023612827063
step: 100, loss: 0.005308025516569614
step: 110, loss: 0.0017291426192969084
step: 120, loss: 0.01208504382520914
step: 130, loss: 0.04489850252866745
step: 140, loss: 0.07966072112321854
step: 150, loss: 0.020449042320251465
step: 160, loss: 0.0010528727434575558

```

```

step: 170, loss: 0.034018248319625854
step: 180, loss: 0.0009106052457354963
step: 190, loss: 0.0007215795340016484
step: 200, loss: 0.004692615941166878
step: 210, loss: 0.019158659502863884
step: 220, loss: 0.017539892345666885
step: 230, loss: 0.009370116516947746
step: 240, loss: 0.03490918129682541
step: 250, loss: 0.012273614294826984
step: 260, loss: 0.012461233884096146
step: 270, loss: 0.008918561972677708
step: 280, loss: 0.007471226621419191
step: 290, loss: 0.13334642350673676
step: 300, loss: 0.004853709600865841
step: 310, loss: 0.0003717902582138777
step: 320, loss: 0.1339874565601349
step: 330, loss: 0.10030553489923477
step: 340, loss: 0.010979698039591312
step: 350, loss: 0.005454869940876961
step: 360, loss: 0.0097414031624794
step: 370, loss: 0.06029887497425079
step: 380, loss: 0.0010281488066539168
step: 390, loss: 0.0016109523130580783
step: 400, loss: 0.04843376576900482
step: 410, loss: 0.004589509218931198
step: 420, loss: 0.0033689034171402454
step: 430, loss: 0.09239795804023743
step: 440, loss: 0.11264947056770325
step: 450, loss: 0.08354994654655457
step: 460, loss: 0.22766190767288208
step: 470, loss: 0.09306012839078903
step: 480, loss: 0.11272281408309937
step: 490, loss: 0.0020758535247296095
step: 500, loss: 0.09462425112724304
step: 510, loss: 0.013783513568341732
step: 520, loss: 0.0024662932846695185
step: 530, loss: 0.09182481467723846
/usr/local/lib/python3.7/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=0.9156398104265402, f1=0.9121813031161473,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0006694080657325685
step: 10, loss: 0.08799700438976288
step: 20, loss: 0.0847768783569336

```

step: 30, loss: 0.0008465080754831433
step: 40, loss: 0.01912321150302887
step: 50, loss: 0.005272553768008947
step: 60, loss: 0.00043224546243436635
step: 70, loss: 0.005195829551666975
step: 80, loss: 0.005484360735863447
step: 90, loss: 0.027921605855226517
step: 100, loss: 0.11614281684160233
step: 110, loss: 0.010088891722261906
step: 120, loss: 0.005826218519359827
step: 130, loss: 0.1323533058166504
step: 140, loss: 0.032076939940452576
step: 150, loss: 0.010544191114604473
step: 160, loss: 0.004344886168837547
step: 170, loss: 0.005307772662490606
step: 180, loss: 0.006589924916625023
step: 190, loss: 0.003623527940362692
step: 200, loss: 0.0007314891554415226
step: 210, loss: 0.01744513213634491
step: 220, loss: 0.0033021203707903624
step: 230, loss: 0.004522337578237057
step: 240, loss: 0.016802135854959488
step: 250, loss: 0.003082931274548173
step: 260, loss: 0.00034006620990112424
step: 270, loss: 0.1717025190591812
step: 280, loss: 0.010088647715747356
step: 290, loss: 0.00073367974255234
step: 300, loss: 0.0024367249570786953
step: 310, loss: 0.013496045023202896
step: 320, loss: 0.0017502796836197376
step: 330, loss: 0.0005428731674328446
step: 340, loss: 0.028255749493837357
step: 350, loss: 0.0007223719730973244
step: 360, loss: 0.0055933184921741486
step: 370, loss: 0.025650236755609512
step: 380, loss: 0.0016168570145964622
step: 390, loss: 0.026640158146619797
step: 400, loss: 0.12957216799259186
step: 410, loss: 0.0016910588601604104
step: 420, loss: 0.0036215982399880886
step: 430, loss: 0.0016898212488740683
step: 440, loss: 0.02370426058769226
step: 450, loss: 0.0029501451645046473
step: 460, loss: 0.007050669752061367
step: 470, loss: 0.0012032013619318604
step: 480, loss: 0.0010656840167939663
step: 490, loss: 0.002080545062199235
step: 500, loss: 0.0003974305291194469

```

step: 510, loss: 0.007140572648495436
step: 520, loss: 0.006157559342682362
step: 530, loss: 0.005808741319924593
/usr/local/lib/python3.7/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=0.9179869524697111, f1=0.9188191881918819,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0033956810366362333
step: 10, loss: 0.008805890567600727
step: 20, loss: 0.001622610492631793
step: 30, loss: 0.0032335345167666674
step: 40, loss: 0.003191061085090041
step: 50, loss: 0.0408569797873497
step: 60, loss: 0.030292915180325508
step: 70, loss: 0.007339250296354294
step: 80, loss: 0.008245952427387238
step: 90, loss: 0.0008941049454733729
step: 100, loss: 0.0012442806037142873
step: 110, loss: 0.0022866916842758656
step: 120, loss: 0.0026555813383311033
step: 130, loss: 0.010252964682877064
step: 140, loss: 0.00023607250477652997
step: 150, loss: 0.0018414624501019716
step: 160, loss: 0.0004129499429836869
step: 170, loss: 0.00016585973207838833
step: 180, loss: 0.0008858799701556563
step: 190, loss: 0.008328258991241455
step: 200, loss: 0.00031627362477593124
step: 210, loss: 0.1316833645105362
step: 220, loss: 0.0009344404097646475
step: 230, loss: 0.0010100039653480053
step: 240, loss: 0.001596852787770331
step: 250, loss: 0.012612021528184414
step: 260, loss: 0.0002410964953014627
step: 270, loss: 0.0010259223636239767
step: 280, loss: 0.0008819462964311242
step: 290, loss: 0.00021817948436364532
step: 300, loss: 0.003135227831080556
step: 310, loss: 0.00012282362149562687
step: 320, loss: 0.0003854755777865648
step: 330, loss: 0.0007413364946842194
step: 340, loss: 0.041635047644376755
step: 350, loss: 0.05031201243400574
step: 360, loss: 0.008572510443627834

```



```
step: 370, loss: 0.005722240079194307
step: 380, loss: 0.06869616359472275
step: 390, loss: 0.004710833076387644
step: 400, loss: 0.0009402727009728551
step: 410, loss: 0.002043582499027252
step: 420, loss: 0.0009896145202219486
step: 430, loss: 0.0004466413811314851
step: 440, loss: 0.0007364929188042879
step: 450, loss: 0.0005073898937553167
step: 460, loss: 0.0009749869350343943
step: 470, loss: 0.033625878393650055
step: 480, loss: 0.06148655340075493
step: 490, loss: 0.004627833608537912
step: 500, loss: 0.0019301821012049913
step: 510, loss: 0.0030717544723302126
step: 520, loss: 0.03718307986855507
step: 530, loss: 0.0020051097963005304
/usr/local/lib/python3.7/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=0.9206927985414768, f1=0.9190172884440401,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0015243538655340672
step: 10, loss: 0.021935539320111275
step: 20, loss: 0.00033364040427841246
step: 30, loss: 0.1490696519613266
step: 40, loss: 0.0004163348057772964
step: 50, loss: 0.004737258423119783
step: 60, loss: 0.0005375686450861394
step: 70, loss: 0.017557473853230476
step: 80, loss: 0.006235751789063215
step: 90, loss: 0.0006075531127862632
step: 100, loss: 0.0001418153551639989
step: 110, loss: 0.0026327345985919237
step: 120, loss: 0.003116433508694172
step: 130, loss: 0.0008631575619801879
step: 140, loss: 0.008319554850459099
step: 150, loss: 0.024342363700270653
step: 160, loss: 0.006006488110870123
step: 170, loss: 0.007557553704828024
step: 180, loss: 0.002085980260744691
step: 190, loss: 0.0035200335551053286
step: 200, loss: 0.0067378925159573555
step: 210, loss: 0.0023958224337548018
step: 220, loss: 0.002016672631725669
```

```

step: 230, loss: 0.0007380839088000357
step: 240, loss: 0.0034803643357008696
step: 250, loss: 0.0358218215405941
step: 260, loss: 0.014104806818068027
step: 270, loss: 0.0031639854423701763
step: 280, loss: 0.001823317026719451
step: 290, loss: 0.03380144014954567
step: 300, loss: 0.00015479527064599097
step: 310, loss: 0.0011821351945400238
step: 320, loss: 0.001139625790528953
step: 330, loss: 0.005863376893103123
step: 340, loss: 0.0006096737924963236
step: 350, loss: 0.020136820152401924
step: 360, loss: 0.005018983036279678
step: 370, loss: 0.0004056965699419379
step: 380, loss: 0.0022599624935537577
step: 390, loss: 0.05368902534246445
step: 400, loss: 0.002450767904520035
step: 410, loss: 0.0006499304436147213
step: 420, loss: 0.0017620909493416548
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 32768.0
step: 430, loss: 0.035330381244421005
step: 440, loss: 0.014528223313391209
step: 450, loss: 0.003125741146504879
step: 460, loss: 0.004184287041425705
step: 470, loss: 0.004488442558795214
step: 480, loss: 0.00262534455396235
step: 490, loss: 0.006643014494329691
step: 500, loss: 0.0017820637440308928
step: 510, loss: 0.0035839430056512356
step: 520, loss: 0.0011852115858346224
step: 530, loss: 0.0007443996146321297
/usr/local/lib/python3.7/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=0.9162790697674419, f1=0.9242843951985227,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0021525609772652388
step: 10, loss: 0.0001997842191485688
step: 20, loss: 0.017205040901899338
step: 30, loss: 0.0005792456795461476
step: 40, loss: 0.0002062551211565733
step: 50, loss: 7.239510887302458e-05
step: 60, loss: 0.0001725324837025255
step: 70, loss: 0.0001371382677461952

```

step: 80, loss: 0.0008196740527637303
step: 90, loss: 6.956225115573034e-05
step: 100, loss: 0.0013100190553814173
step: 110, loss: 7.26299622328952e-05
step: 120, loss: 0.0014849354047328234
step: 130, loss: 0.11121488362550735
step: 140, loss: 0.001695145620033145
step: 150, loss: 6.136237789178267e-05
step: 160, loss: 0.0014976896345615387
step: 170, loss: 0.01946667768061161
step: 180, loss: 9.834948286879808e-05
step: 190, loss: 0.11007821559906006
step: 200, loss: 0.013028758578002453
step: 210, loss: 8.327340765390545e-05
step: 220, loss: 0.012062161229550838
step: 230, loss: 0.007641766685992479
step: 240, loss: 0.0005839609657414258
step: 250, loss: 0.0017550501506775618
step: 260, loss: 0.003742692992091179
step: 270, loss: 0.0013317072298377752
step: 280, loss: 0.00010087835835292935
step: 290, loss: 0.02664172649383545
step: 300, loss: 6.699084042338654e-05
step: 310, loss: 0.008692469447851181
step: 320, loss: 0.0007810689858160913
step: 330, loss: 0.000710830558091402
step: 340, loss: 0.0006137603195384145
Gradient overflow. Skipping step, loss scaler 0 reducing loss scale to 16384.0
step: 350, loss: 0.01568085141479969
step: 360, loss: 4.780217204825021e-05
step: 370, loss: 0.00044553185580298305
step: 380, loss: 6.348468014039099e-05
step: 390, loss: 0.00032833300065249205
step: 400, loss: 0.0009033515816554427
step: 410, loss: 0.00021522477618418634
step: 420, loss: 0.0001234511291841045
step: 430, loss: 0.0002709035179577768
step: 440, loss: 0.0009226789698004723
step: 450, loss: 0.0009364402503706515
step: 460, loss: 0.0009379658149555326
step: 470, loss: 0.016788868233561516
step: 480, loss: 0.01588280498981476
step: 490, loss: 0.004245594143867493
step: 500, loss: 0.0025770440697669983
step: 510, loss: 0.00022311195789370686
step: 520, loss: 0.00039774354081600904
step: 530, loss: 0.00019451286061666906
/usr/local/lib/python3.7/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=0.9164743885556068, f1=0.9202566452795601,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00027291924925521016
step: 10, loss: 0.00019193108892068267
step: 20, loss: 0.00011109344632131979
step: 30, loss: 4.44293582404498e-05
step: 40, loss: 0.0002961991704069078
step: 50, loss: 0.0019619239028543234
step: 60, loss: 0.0068040881305933
step: 70, loss: 0.00014185119653120637
step: 80, loss: 0.0022264255676418543
step: 90, loss: 0.0005015474162064493
step: 100, loss: 0.0007721108268015087
step: 110, loss: 0.001803336781449616
step: 120, loss: 0.0011198935098946095
step: 130, loss: 0.0002930292976088822
step: 140, loss: 0.005612341687083244
step: 150, loss: 0.25187948346138
step: 160, loss: 0.0007090115104801953
step: 170, loss: 0.012580416165292263
step: 180, loss: 0.001482407096773386
step: 190, loss: 0.002418265212327242
step: 200, loss: 0.00044505661935545504
step: 210, loss: 0.00018751119205262512
step: 220, loss: 0.0034412285313010216
step: 230, loss: 0.004826510325074196
step: 240, loss: 0.0020293141715228558
step: 250, loss: 0.00012497109128162265
step: 260, loss: 0.001707238145172596
step: 270, loss: 0.0025163600221276283
step: 280, loss: 0.0006627935217693448
step: 290, loss: 0.0003648278070613742
step: 300, loss: 0.00013216091610956937
step: 310, loss: 0.0001964621478691697
step: 320, loss: 0.0014955932274460793
step: 330, loss: 0.000331792893121019
step: 340, loss: 0.0008968464098870754
step: 350, loss: 0.0006174652953632176
step: 360, loss: 0.005673261359333992
step: 370, loss: 0.0007976102060638368
step: 380, loss: 0.005921169184148312
step: 390, loss: 0.002032096032053232
step: 400, loss: 0.0006449001957662404
```

```

step: 410, loss: 0.002782230731099844
step: 420, loss: 0.0006281029782257974
step: 430, loss: 0.00010919332999037579
step: 440, loss: 0.00020162694272585213
step: 450, loss: 0.024738803505897522
step: 460, loss: 0.008463197387754917
step: 470, loss: 0.000278828723821789
step: 480, loss: 0.0004085532855242491
step: 490, loss: 0.0009796854574233294
step: 500, loss: 0.033432673662900925
step: 510, loss: 5.948938269284554e-05
step: 520, loss: 0.0004933146410621703
step: 530, loss: 0.0022473011631518602
/usr/local/lib/python3.7/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=0.9172033118675251, f1=0.9155963302752294,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0004086287517566234
step: 10, loss: 0.01301321666687727
step: 20, loss: 0.001207307679578662
step: 30, loss: 0.0012507513165473938
step: 40, loss: 0.0006665317341685295
step: 50, loss: 0.0022278791293501854
step: 60, loss: 0.011998087167739868
step: 70, loss: 0.0727248266339302
step: 80, loss: 0.002681497484445572
step: 90, loss: 0.00044941456872038543
step: 100, loss: 0.00036087504122406244
step: 110, loss: 0.0011035241186618805
step: 120, loss: 0.002143182558938861
step: 130, loss: 0.0004358085279818624
step: 140, loss: 0.07179993391036987
step: 150, loss: 0.0001383048656862229
step: 160, loss: 0.0002239429159089923
step: 170, loss: 0.00030307675478979945
step: 180, loss: 0.0011674920096993446
step: 190, loss: 0.0006041631568223238
step: 200, loss: 0.010846195742487907
step: 210, loss: 0.0002193587861256674
step: 220, loss: 0.0037538865581154823
step: 230, loss: 0.00039608628139831126
step: 240, loss: 0.0003263847902417183
step: 250, loss: 6.14236414548941e-05
step: 260, loss: 4.122109749005176e-05

```

```

step: 270, loss: 0.0015926264459267259
step: 280, loss: 0.000324654538417235
step: 290, loss: 0.007876588962972164
step: 300, loss: 0.0012890022480860353
step: 310, loss: 0.00040184918907471
step: 320, loss: 4.7400255425600335e-05
step: 330, loss: 8.432790491497144e-05
step: 340, loss: 0.0001989943120861426
step: 350, loss: 0.0008296450832858682
step: 360, loss: 0.00012427335605025291
step: 370, loss: 0.00032122741686180234
step: 380, loss: 0.00022537099721375853
step: 390, loss: 0.00011750697740353644
step: 400, loss: 0.0001074769752449356
step: 410, loss: 0.0060130637139081955
step: 420, loss: 0.0004319415311329067
step: 430, loss: 7.387444202322513e-05
step: 440, loss: 0.000959812372457236
step: 450, loss: 0.00010046276292996481
step: 460, loss: 0.13782931864261627
step: 470, loss: 0.00022170953161548823
step: 480, loss: 0.002544717863202095
step: 490, loss: 0.002089774701744318
step: 500, loss: 0.00017467650468461215
step: 510, loss: 0.03620464727282524
step: 520, loss: 0.00014531788474414498
step: 530, loss: 0.0437837578356266
/usr/local/lib/python3.7/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=0.9222846441947565, f1=0.9133271202236719,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
  warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.0014268382219597697
step: 10, loss: 0.00029826484387740493
step: 20, loss: 0.00016141746891662478
step: 30, loss: 0.043890777975320816
step: 40, loss: 0.00019799704023171216
step: 50, loss: 0.19533772766590118
step: 60, loss: 0.001383391791023314
step: 70, loss: 0.0014540111878886819
step: 80, loss: 0.00011512244236655533
step: 90, loss: 0.0003890385269187391
step: 100, loss: 0.01821916177868843
step: 110, loss: 7.366965292021632e-05
step: 120, loss: 0.11653763055801392

```

```
step: 130, loss: 7.041706703603268e-05
step: 140, loss: 9.579704055795446e-05
step: 150, loss: 2.1535497580771334e-05
step: 160, loss: 6.49096182314679e-05
step: 170, loss: 0.030184239149093628
step: 180, loss: 0.00013230566401034594
step: 190, loss: 0.00013988104183226824
step: 200, loss: 0.00040043049375526607
step: 210, loss: 0.0018766806460916996
step: 220, loss: 8.192887617042288e-05
step: 230, loss: 4.5742886868538335e-05
step: 240, loss: 0.0012015430256724358
step: 250, loss: 0.00034254734055139124
step: 260, loss: 4.6560351620428264e-05
step: 270, loss: 0.0010338915744796395
step: 280, loss: 0.00031812043744139373
step: 290, loss: 0.0009428983321413398
step: 300, loss: 0.05172441154718399
step: 310, loss: 0.00010797494906000793
step: 320, loss: 0.00011194371472811326
step: 330, loss: 0.0001569471787661314
step: 340, loss: 0.00026301087928004563
step: 350, loss: 0.0007987364660948515
step: 360, loss: 0.005952878855168819
step: 370, loss: 0.012306177988648415
step: 380, loss: 0.004447312094271183
step: 390, loss: 0.0011310795089229941
step: 400, loss: 0.030268950387835503
step: 410, loss: 7.199820538517088e-05
step: 420, loss: 0.006682489067316055
step: 430, loss: 0.004289600532501936
step: 440, loss: 0.07079697400331497
step: 450, loss: 0.0234861858189106
step: 460, loss: 0.00035255946568213403
step: 470, loss: 0.00150804512668401
step: 480, loss: 0.000640198471955955
step: 490, loss: 0.00030720746144652367
step: 500, loss: 3.2258740247925743e-05
step: 510, loss: 0.028986793011426926
step: 520, loss: 4.450900451047346e-05
step: 530, loss: 0.00013659559772349894
/usr/local/lib/python3.7/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=0.9211147850732168, f1=0.9201127819548872,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
```

```
warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.004196262918412685
step: 10, loss: 0.0022130985744297504
step: 20, loss: 0.004879768006503582
step: 30, loss: 0.00040026201168075204
step: 40, loss: 0.0001554218033561483
step: 50, loss: 0.003561090910807252
step: 60, loss: 9.703898831503466e-05
step: 70, loss: 2.6410316422698088e-05
step: 80, loss: 6.66006890241988e-05
step: 90, loss: 0.0033032670617103577
step: 100, loss: 0.0002315793390152976
step: 110, loss: 9.371408668812364e-05
step: 120, loss: 0.00018974865088239312
step: 130, loss: 1.6655469153192826e-05
step: 140, loss: 0.00010915447637671605
step: 150, loss: 0.001132722944021225
step: 160, loss: 7.220112456707284e-05
step: 170, loss: 0.0006825723103247583
step: 180, loss: 9.828728070715442e-05
step: 190, loss: 0.0002075399534078315
step: 200, loss: 0.0002408798027317971
step: 210, loss: 7.989844743860886e-05
step: 220, loss: 0.0001278335985261947
step: 230, loss: 8.819713548291475e-05
step: 240, loss: 0.00017561562708579004
step: 250, loss: 4.607881419360638e-05
step: 260, loss: 8.677581354277208e-05
step: 270, loss: 3.9751270378474146e-05
step: 280, loss: 0.00027854906511493027
step: 290, loss: 0.0005282895872369409
step: 300, loss: 5.213259646552615e-05
step: 310, loss: 0.0020554526709020138
step: 320, loss: 0.010212940163910389
step: 330, loss: 0.01681152544915676
step: 340, loss: 0.00018630531849339604
step: 350, loss: 3.647457924671471e-05
step: 360, loss: 3.373375875526108e-05
step: 370, loss: 7.047411781968549e-05
step: 380, loss: 0.00018511329835746437
step: 390, loss: 0.0003141034103464335
step: 400, loss: 0.0009638765477575362
step: 410, loss: 0.0018941254820674658
step: 420, loss: 9.431014768779278e-05
step: 430, loss: 0.0026656801346689463
step: 440, loss: 0.00031635872437618673
step: 450, loss: 0.0001567101280670613
step: 460, loss: 0.00032226971234194934
```



```

step: 470, loss: 0.0008274338906630874
step: 480, loss: 8.476414950564504e-05
step: 490, loss: 0.0006805384182371199
step: 500, loss: 3.08065464196261e-05
step: 510, loss: 0.00011966231249971315
step: 520, loss: 0.0009427727782167494
step: 530, loss: 0.0016200292157009244
/usr/local/lib/python3.7/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=0.918918918918919, f1=0.919815668202765,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 0.00016592119936831295
step: 10, loss: 1.6230851542786695e-05
step: 20, loss: 0.0003213815507479012
step: 30, loss: 1.7147209291579202e-05
step: 40, loss: 0.03551679477095604
step: 50, loss: 0.0045738425105810165
step: 60, loss: 1.2270999832253437e-05
step: 70, loss: 5.6733813835307956e-05
step: 80, loss: 4.6462479076581076e-05
step: 90, loss: 0.0019668268505483866
step: 100, loss: 0.0006074834964238107
step: 110, loss: 7.35408320906572e-05
step: 120, loss: 0.0002833390317391604
step: 130, loss: 0.007994208484888077
step: 140, loss: 0.000162361073307693
step: 150, loss: 0.00024154283164534718
step: 160, loss: 8.194385009119287e-05
step: 170, loss: 0.0032390733249485493
step: 180, loss: 0.03219957277178764
step: 190, loss: 0.00034632149618119
step: 200, loss: 0.0033102366141974926
step: 210, loss: 0.0001889999839477241
step: 220, loss: 0.002996159950271249
step: 230, loss: 5.9321853768778965e-05
step: 240, loss: 0.014077380299568176
step: 250, loss: 4.0136750612873584e-05
step: 260, loss: 3.915916022378951e-05
step: 270, loss: 9.373675129609182e-05
step: 280, loss: 1.6770927686593495e-05
step: 290, loss: 0.0001812939008232206
step: 300, loss: 0.0004922301159240305
step: 310, loss: 0.022495390847325325
step: 320, loss: 0.00020793756993953139

```

```

step: 330, loss: 0.0012545189820230007
step: 340, loss: 0.0001407687523169443
step: 350, loss: 0.0014742386993020773
step: 360, loss: 0.012394745834171772
step: 370, loss: 3.211012881365605e-05
step: 380, loss: 0.0005555456154979765
step: 390, loss: 0.003884383710101247
step: 400, loss: 0.00011171957885380834
step: 410, loss: 1.6916272215894423e-05
step: 420, loss: 7.483080844394863e-05
step: 430, loss: 0.0001581687101861462
step: 440, loss: 0.0024400034453719854
step: 450, loss: 9.980182949220762e-05
step: 460, loss: 4.942353552905843e-05
step: 470, loss: 0.0002464594435878098
step: 480, loss: 0.00012375987716950476
step: 490, loss: 0.00010846122313523665
step: 500, loss: 9.97531897155568e-05
step: 510, loss: 5.5101227189879864e-05
step: 520, loss: 3.872405068250373e-05
step: 530, loss: 0.00015742173127364367
/usr/local/lib/python3.7/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=0.9209171736078615, f1=0.9215048769159313,
best_f1=0.9215777262180975
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
step: 0, loss: 2.019068051595241e-05
step: 10, loss: 0.00015174297732301056
step: 20, loss: 2.9819242627127096e-05
step: 30, loss: 8.307072130264714e-05
step: 40, loss: 0.02194751799106598
step: 50, loss: 0.03407355770468712
step: 60, loss: 1.989269549085293e-05
step: 70, loss: 0.0002259107568534091
step: 80, loss: 0.0001035901004797779
step: 90, loss: 2.411613240838051e-05
step: 100, loss: 0.0007764214533381164
step: 110, loss: 0.00022583885584026575
step: 120, loss: 6.842219590907916e-05
step: 130, loss: 0.010267670266330242
step: 140, loss: 6.922987085999921e-05
step: 150, loss: 2.5330979042337276e-05
step: 160, loss: 0.0001748408831190318
step: 170, loss: 1.802617589419242e-05
step: 180, loss: 0.00011215040285605937

```

```

step: 190, loss: 3.998599640908651e-05
step: 200, loss: 5.815970871481113e-05
step: 210, loss: 0.0007958183414302766
step: 220, loss: 0.0006948591908439994
step: 230, loss: 0.001014119596220553
step: 240, loss: 0.00010182345431530848
step: 250, loss: 0.000220369067392312
step: 260, loss: 8.643844921607524e-05
step: 270, loss: 1.3757147826254368e-05
step: 280, loss: 1.3083067642583046e-05
step: 290, loss: 1.3165048585506156e-05
step: 300, loss: 4.5721848437096924e-05
step: 310, loss: 0.0001700740249361843
step: 320, loss: 2.635735654621385e-05
step: 330, loss: 2.3846889234846458e-05
step: 340, loss: 2.8111679057474248e-05
step: 350, loss: 0.0002890195173677057
step: 360, loss: 0.00023661630984861404
step: 370, loss: 5.4895092034712434e-05
step: 380, loss: 4.2946809116983786e-05
step: 390, loss: 0.0085041094571352
step: 400, loss: 0.0001134208869189024
step: 410, loss: 3.98781594412867e-05
step: 420, loss: 0.00016078643966466188
step: 430, loss: 0.0009365401929244399
step: 440, loss: 6.489369843620807e-05
step: 450, loss: 1.4174458556226455e-05
step: 460, loss: 1.2375242476991843e-05
step: 470, loss: 0.0002497888926882297
step: 480, loss: 1.8592174455989152e-05
step: 490, loss: 0.0001104616021621041
step: 500, loss: 2.041393963736482e-05
step: 510, loss: 3.0944040190661326e-05
step: 520, loss: 0.002612848998978734
step: 530, loss: 1.348166279058205e-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=0.9209302325581395, f1=0.917397323488694,
best_f1=0.9215777262180975
mv: cannot stat '*_dev.pt': No such file or directory
Selected optimization level 02:  FP16 training with FP32 batchnorm and FP32
master weights.

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

```

```

patch_torch_functions : False
keep_batchnorm_fp32   : True
master_weights        : True
loss_scale             : dynamic
Processing user overrides (additional kwargs that are not None)...
After processing overrides, optimization options are:
enabled                : True
opt_level              : 02
cast_model_type        : torch.float16
patch_torch_functions : False
keep_batchnorm_fp32    : True
master_weights         : True
loss_scale             : dynamic
Warning: multi_tensor_applier fused unscale kernel is unavailable, possibly
because apex was installed without --cuda_ext --cpp_ext. Using Python fallback.
Original ImportError was: ModuleNotFoundError("No module named 'amp_C'")
/usr/local/lib/python3.7/dist-packages/apex/amp/_initialize.py:25: UserWarning:
An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
0it [00:00, ?it/s]/usr/local/lib/python3.7/dist-
packages/apex/amp/_initialize.py:25: UserWarning: An input tensor was not cuda.
    warnings.warn("An input tensor was not cuda.")
5742it [00:22, 253.94it/s]
load_f1 = 0.9222222222222223
real_f1 = 0.9223616922361693
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, 257.81it/s]

```