CA4P-483

Public Repository: https://github.com/zacharykzhao/CA4P-483

1. Download the repository

Run the following script in your terminal (please make sure ***git***(https://git-scm.com/book/en/v2/Getting-Started-Installing-Git) has been correctly installed):

```
git clone https://github.com/zacharykzhao/CA4P-483
```

```
[$ git clone https://github.com/zacharykzhao/CA4P-483
Cloning into 'CA4P-483'...
remote: Enumerating objects: 1696, done.
remote: Counting objects: 100% (40/40), done.
remote: Compressing objects: 100% (37/37), done.
remote: Total 1696 (delta 22), reused 6 (delta 2), pack-reused 1656
Receiving objects: 100% (1696/1696), 42.97 MiB | 10.65 MiB/s, done.
Resolving deltas: 100% (713/713), done.
Updating files: 100% (1137/1137), done.
```

2. Setup

Execute the following script in command line to create the python interpreter environment to reproduce the results:

```
conda create -n ca4p python=3.6
conda activate ca4p
pip install -r ./CA4P-483/requirements.txt
```

3. Get information in Table 1

- 3.1 Run the following script in command line to reproduce statistic information in Table 1:

```
python ./CA4P-483/DataSetStatistics/Table_1.py
```

```
[$ python ./CA4P-483/DataSetStatistics/Table_1.py
# doc   483
------------------------
# sentences     11565
------------------------
# sentences with ann    3385
------------------------
Avg sentences len       79.06
------------------------
Type            Num             Train           Dev             Test            Avg len
------------------------
data            21241           18925           2521            2331            4.68
collect                 5134            4133            576             528             2.03
share           4976            3989            533             505             2.10
handler                 8424            6085            815             782             2.49
condition               4917            5477            716             713             14.41
subjects                3202            2776            360             350             4.29
purpose                 4683            6442            860             867             19.24
------------------------
Total           52577           47827           6381            6076
```

- Please notice that the "handler" in the script corresponds to "Controller" in the paper, and the "subjects" in the script conrresponds to "Receiver" in the paper.

- 3.2 To obatian the **Kappa** agreements for each components in Table 1, please refer to the form: "./CA4P-483/Manual Kappa Evaluation 0611.xlsx".

In the form, row 1-9 denotes the manual check results of the 1st annotator, row 12-20 refers to the manual check results of the 2nd annotator, and row 25-33 gives the **Kappa** aggreements based on the manual check results, i.e., row 1-20.

4. Reproduce results in Table 3 & 4

- 4.1 For CRF, HMM, Bi-LSTM, Bi-LSTM-CRF:
  - 4.1.1 Execute the following script in command line to create the python interpreter virtual environment to reproduce the results. Besides, please make sure that the pytorch (GPU version is strongly recommended) is correctly installed in your environment.

```
conda create -n CHBB python=3.6
conda activate CHBB
pip install -r ./CA4P-483/baselines/named_entity_recognition-
master/requirement.txt
```

  -

```
[$ pip install -r ./CA4P-483/baselines/named_entity_recognition-master/requirement.txt                                ]
Collecting numpy==1.16.2
  Downloading numpy-1.16.2-cp36-cp36m-macosx_10_6_intel.macosx_10_9_intel.macosx_10_9_x86_64.macosx_10_10_intel.macosx_10_10_x86_64.w
hl (13.9 MB)
     |████████████████████████████████| 13.9 MB 490 kB/s
Collecting python-crfsuite==0.9.6
  Downloading python_crfsuite-0.9.6-cp36-cp36m-macosx_10_13_x86_64.whl (192 kB)
     |████████████████████████████████| 192 kB 52.7 MB/s
Collecting six==1.12.0
  Downloading six-1.12.0-py2.py3-none-any.whl (10 kB)
Collecting sklearn-crfsuite==0.3.6
  Downloading sklearn_crfsuite-0.3.6-py2.py3-none-any.whl (12 kB)
Collecting tabulate==0.8.3
  Downloading tabulate-0.8.3.tar.gz (46 kB)
     |████████████████████████████████| 46 kB 15.3 MB/s
Collecting torch==1.0.1.post2
  Downloading torch-1.0.1.post2-cp36-none-macosx_10_7_x86_64.whl (27.2 MB)
     |████████████████████████████████| 27.2 MB 417 kB/s
Collecting tqdm==4.31.1
  Downloading tqdm-4.31.1-py2.py3-none-any.whl (48 kB)
     |████████████████████████████████| 48 kB 14.5 MB/s
Building wheels for collected packages: tabulate
  Building wheel for tabulate (setup.py) ... done
  Created wheel for tabulate: filename=tabulate-0.8.3-py3-none-any.whl size=23391 sha256=9d6187ac4c5664d1f4f1924f634a1887b592dbcc893c
e13ed495ca51167af92a
  Stored in directory: /Users/zhaokaifa_imac/Library/Caches/pip/wheels/20/2a/38/bbf076580de21676f282b7a40156184b9ef5bc8ebca4dbddb1
Successfully built tabulate
Installing collected packages: tqdm, tabulate, six, python-crfsuite, torch, sklearn-crfsuite, numpy
Successfully installed numpy-1.16.2 python-crfsuite-0.9.6 six-1.12.0 sklearn-crfsuite-0.3.6 tabulate-0.8.3 torch-1.0.1.post2 tqdm-4.3
1.1
```

- 4.1.2 Execute the following script in command line to reproduce the results:

```
cd CA4P-483/baselines/named_entity_recognition-master
python CSP3_main.py
```

  - CRF Results:

```
正在训练评估CRF模型...
              precision    recall  f1-score   support

       E-data     0.7675    0.7044    0.7346      2324
   I-collect     0.7500    0.3158    0.4444        19
   E-condition     0.5903    0.4586    0.5162       713
   E-collect     0.7505    0.7419    0.7462       527
   I-condition     0.6545    0.5062    0.5709      5841
   B-handler     0.6388    0.4448    0.5245       843
     I-share     0.4048    0.5397    0.4626        63
   B-condition     0.5415    0.4202    0.4732       714
   I-handler     0.7442    0.2739    0.4004       701
     B-share     0.7285    0.6376    0.6800       505
      I-data     0.7547    0.7031    0.7280      6335
   B-subjects     0.6389    0.4574    0.5331       352
   E-purpose     0.7723    0.5600    0.6493       866
   B-purpose     0.5764    0.4166    0.4836       869
   B-collect     0.7505    0.7405    0.7455       528
     E-share     0.7738    0.6772    0.7223       505
   I-subjects     0.5698    0.4093    0.4764       838
   E-subjects     0.7500    0.5478    0.6332       345
   E-handler     0.6495    0.4547    0.5349       750
      B-data     0.7345    0.6745    0.7032      2329
           O     0.8930    0.9423    0.9170    115128
   I-purpose     0.7756    0.5662    0.6546      7179
   avg/total     0.8535    0.8614    0.8546    148274
```

- HMM Results:

```
正在训练评估HMM模型...
              precision    recall  f1-score   support

      E-data     0.4129    0.7285    0.5271      2324
   I-collect     0.0000    0.0000    0.0000        19
   E-condition     0.2242    0.4320    0.2952       713
   E-collect     0.3029    0.7495    0.4315       527
   I-condition     0.2962    0.5316    0.3804      5841
   B-handler     0.2393    0.6251    0.3461       843
     I-share     0.1493    0.8889    0.2557        63
   B-condition     0.2147    0.4132    0.2826       714
   I-handler     0.1403    0.2140    0.1695       701
     B-share     0.2244    0.8436    0.3546       505
      I-data     0.4148    0.7214    0.5267      6335
   B-subjects     0.1618    0.5142    0.2461       352
   E-purpose     0.3431    0.4619    0.3937       866
   B-purpose     0.2356    0.3165    0.2701       869
   B-collect     0.3037    0.7500    0.4323       528
     E-share     0.2397    0.9010    0.3787       505
   I-subjects     0.1955    0.7160    0.3071       838
   E-subjects     0.2091    0.6783    0.3197       345
   E-handler     0.2393    0.7000    0.3567       750
      B-data     0.3632    0.6393    0.4632      2329
           O     0.9016    0.6684    0.7676    115128
   I-purpose     0.4009    0.5891    0.4771      7179
   avg/total     0.7755    0.6559    0.6921    148274
```

- BiLSTM results:
  -

```
正在训练评估双向LSTM模型...
Epoch 1, step/total_step: 5/230 2.17% Loss:3.1398
Epoch 1, step/total_step: 10/230 4.35% Loss:2.7698
Epoch 1, step/total_step: 15/230 6.52% Loss:2.1681
Epoch 1, step/total_step: 20/230 8.70% Loss:1.4350
Epoch 1, step/total_step: 25/230 10.87% Loss:1.3683
Epoch 1, step/total_step: 30/230 13.04% Loss:1.4219
Epoch 1, step/total_step: 35/230 15.22% Loss:1.4341
Epoch 1, step/total_step: 40/230 17.39% Loss:1.2851
Epoch 1, step/total_step: 45/230 19.57% Loss:1.2601
Epoch 1, step/total_step: 50/230 21.74% Loss:1.0465
Epoch 1, step/total_step: 55/230 23.91% Loss:1.1540
Epoch 1, step/total_step: 60/230 26.09% Loss:1.0149
Epoch 1, step/total_step: 65/230 28.26% Loss:1.0685
Epoch 1, step/total_step: 70/230 30.43% Loss:1.0279
Epoch 1, step/total_step: 75/230 32.61% Loss:0.9939
Epoch 1, step/total_step: 80/230 34.78% Loss:0.9363
Epoch 1, step/total_step: 85/230 36.96% Loss:0.9799
Epoch 1, step/total_step: 90/230 39.13% Loss:1.0047
```

```
训练完毕,共用时743秒.
评估bilstm模型中...
/home/zachary/anaconda3/envs/CHBB/lib/python3.7/site-packag
rt of single contiguous chunk of memory. This means they ne
. To compact weights again call flatten_parameters(). (Trig
  self.num_layers, self.dropout, self.training, self.bidire
              precision    recall  f1-score   support

   E-handler     0.8106    0.5460    0.6525       674
   E-purpose     0.6569    0.6071    0.6310       845
   B-handler     0.7828    0.5290    0.6314       845
   B-subjects     0.7086    0.3523    0.4706       352
      I-data     0.8256    0.6045    0.6980      6443
I-condition     0.6025    0.5348    0.5666      5896
      B-data     0.8374    0.5723    0.6799      2331
I-subjects     0.6988    0.3458    0.4627       986
I-purpose     0.6600    0.5888    0.6224      7209
           O     0.8899    0.9476    0.9179    115225
   B-collect     0.8060    0.6125    0.6960       529
B-condition     0.5341    0.3950    0.4541       714
     I-share     0.6190    0.2047    0.3077       127
E-subjects     0.8205    0.4873    0.6115       197
     E-share     0.8174    0.6205    0.7054       440
     B-share     0.7665    0.5968    0.6711       506
B-purpose     0.5927    0.3563    0.4451       870
   E-collect     0.7872    0.5717    0.6624       453
   I-collect     0.7143    0.1596    0.2609        94
   I-handler     0.6817    0.2694    0.3862       787
E-condition     0.5484    0.4392    0.4878       658
      E-data     0.8415    0.6565    0.7376      2224
avg/total     0.8512    0.8588    0.8509    148405
```

- BiLSTM-CRF

```
正在训练评估Bi-LSTM+CRF模型...
Epoch 1, step/total_step: 5/230 2.17% Loss:1278.6317
Epoch 1, step/total_step: 10/230 4.35% Loss:682.1607
Epoch 1, step/total_step: 15/230 6.52% Loss:424.7988
Epoch 1, step/total_step: 20/230 8.70% Loss:252.8031
Epoch 1, step/total_step: 25/230 10.87% Loss:205.2397
Epoch 1, step/total_step: 30/230 13.04% Loss:189.0960
Epoch 1, step/total_step: 35/230 15.22% Loss:173.1340
Epoch 1, step/total_step: 40/230 17.39% Loss:146.2633
Epoch 1, step/total_step: 45/230 19.57% Loss:132.0924
Epoch 1, step/total_step: 50/230 21.74% Loss:105.0311
Epoch 1, step/total_step: 55/230 23.91% Loss:111.3476
Epoch 1, step/total_step: 60/230 26.09% Loss:91.2688
Epoch 1, step/total_step: 65/230 28.26% Loss:92.0068
Epoch 1, step/total_step: 70/230 30.43% Loss:84.3765
Epoch 1, step/total_step: 75/230 32.61% Loss:80.1204
Epoch 1, step/total_step: 80/230 34.78% Loss:71.9153
Epoch 1, step/total_step: 85/230 36.96% Loss:73.2391
Epoch 1, step/total_step: 90/230 39.13% Loss:75.0921
Epoch 1, step/total_step: 95/230 41.30% Loss:55.9308
Epoch 1, step/total_step: 100/230 43.48% Loss:57.0105
Epoch 1, step/total_step: 105/230 45.65% Loss:54.7476
Epoch 1, step/total_step: 110/230 47.83% Loss:47.2397
Epoch 1, step/total_step: 115/230 50.00% Loss:51.0544
Epoch 1, step/total_step: 120/230 52.17% Loss:42.3588
Epoch 1, step/total_step: 125/230 54.35% Loss:36.9988
Epoch 1, step/total_step: 130/230 56.52% Loss:36.3080
Epoch 1, step/total_step: 135/230 58.70% Loss:45.5186
Epoch 1, step/total_step: 140/230 60.87% Loss:36.3204
Epoch 1, step/total_step: 145/230 63.04% Loss:32.5107
Epoch 1, step/total_step: 150/230 65.22% Loss:31.9330
Epoch 1, step/total_step: 155/230 67.39% Loss:24.5726
Epoch 1, step/total_step: 160/230 69.57% Loss:31.1649
Epoch 1, step/total_step: 165/230 71.74% Loss:23.7172
Epoch 1, step/total_step: 170/230 73.91% Loss:25.1459
Epoch 1, step/total_step: 175/230 76.09% Loss:23.0954
Epoch 1, step/total_step: 180/230 78.26% Loss:25.1449
```

```
评估bilstm_crf模型中...
              precision    recall  f1-score   support

   E-handler     0.8394    0.6513    0.7335       674
   E-purpose     0.6857    0.6686    0.6771       845
   B-handler     0.8336    0.6343    0.7204       845
   B-subjects     0.6553    0.5455    0.5953       352
      I-data     0.8125    0.6665    0.7323      6443
  I-condition     0.5809    0.5746    0.5778       589
      B-data     0.8000    0.6589    0.7227      2331
  I-subjects     0.6140    0.5517    0.5812       986
  I-purpose     0.6549    0.7002    0.6768      7209
           O     0.9074    0.9300    0.9186    115225
  B-collect     0.8341    0.7032    0.7631       529
  B-condition     0.4824    0.4608    0.4713       714
    I-share     0.5091    0.2205    0.3077       127
  E-subjects     0.7800    0.5939    0.6744       197
    E-share     0.7980    0.7091    0.7509       440
    B-share     0.7836    0.6798    0.7280       506
  B-purpose     0.5089    0.4908    0.4997       870
  E-collect     0.7990    0.7108    0.7523       453
  I-collect     0.6170    0.3085    0.4113        94
  I-handler     0.7249    0.4320    0.5414       787
  E-condition     0.5182    0.4970    0.5074       658
     E-data     0.8361    0.6951    0.7591      2224
avg/total     0.8617    0.8638    0.8616    148405
```

- Confusion Matrix for BiLSTM-CRF (Corresponds to Figure2)

```
Confusion Matrix:
       E-handler E-purpose B-handler B-subjects  I-data I-condition  B-data I-subjects I-purpose      O B-collect B-condition I-share E-subjects E-share B-share B-purpose E-collect I-collect I-handler E-condition  E-data
E-handler     439         0         0         0       0         2       0         1         0    220         0           0       2         0       0       1         0       7         2         0           0
E-purpose       0       565         0         0       0         9       0         0         7    256         0           0       0         0       0       0         0       0         7         1           0
B-handler       1         0       536         3       0         3       0         0         1    291         1           5       0         0       0       0         0       4         0         0           0
B-subjects      0         0         2       192       0         3       0        31         2    119         0           3       0         0       0       0         0       0         0         0           0
 I-data         0         0         0         0    4294         8      65         0        23   2029         0           0       0         0       0       0         0       0         0        24           0
I-condition     0         4         0         0      11      3388       0         0       126   2269         0          51       0         0       2       2         5       0         0         0          38           0
 B-data         0         0         0         0      68         4    1536         0         2    716         1           1       0         0       0       3         0       0         0         0           0           0
I-subjects      0         0         0        17       4        18       1       544         3    372         0           0       0         4       0       0         0       0        23         0           0           0
I-purpose       0         2         0         0       8        81       2         0      5048   2026         0           1       0         0       0      40         1       0         0         0           0           0
        O      71       242        97        81     863      2166     313       289      2369 107157        67         285      19        27      57      77       360      58         9        95         248         275
B-collect       0         0         0         0       0         0       0         0         1    155       372           0       0         0       0       1         0       0         0         0           0           0
B-condition     0         0         0         0       1        68       1         0         2    310         0         329       0         0       0       0         3       0         0         0           0           0           0
I-share         0         0         0         0       0         0       0         0         0     66         0           0      28         0      18      15         0       0         0         0           0           0
E-subjects      2         1         0         0       0         2       0        16         1     53         0           0       0       117       0       0         0       0         0         0           4           1
E-share         0         0         0         0       1         4       0         0         1    116         0           0       5         0     312       0         0       0         0         0           1           0
B-share         0         0         0         0       0         5       1         0         0    151         0           0       3         0       1     344         1       0         0         0           0           0
B-purpose       0         0         0         0       0         8       1         0       118    308         1           7       0         0       0       0       427       0         0         0           0           0
E-collect       0         0         0         0       0         0       0         0         0    122         0           0       0         0       0       0         0     322         9         0           0           0
I-collect       0         0         0         0       0         0       0         0         1     41         3           0       0         0       1       0         0      19        29         0           0           0
I-handler      10         0         8         0       0        15       0         5         1    406         1           0       0         0       0       0         0       1         0       340           0           0
E-condition     0         6         0         0       0        47       0         0         1    275         0           0       0         0       0       0         0       0         0         0         327           2
 E-data         0         4         0         0      35         1       0         0         1    632         0           0       0         0       0       0         0       1         0         0         4        1546
Traceback (most recent call last):
  File "CSP3_main.py", line 112, in <module>
```

- **It is wothy noticing that HMM, CRF, BiLSTM and BiLSTM-CRF calculate the metric for each class/components with B, I, E seperately. We average them to obtain the results in Table 3.**

- 4.2 For BERT-BiLSTM-CRF:
  - 4.2.1 Execute the following script in command line to create the python interpreter virtual environment to reproduce the results:

```
conda create -n BBC python=3.9
conda activate BBC
cd ../BERT-BiLSTM-CRF-NER-tf-master
pip install -r requirement.txt
python3 setup.py install
```

```
$ pip install -r requirement.txt

Collecting GPUtil>=1.3.0 (from -r requirement.txt (line 4))
  Downloading GPUtil-1.4.0.tar.gz (5.5 kB)
  Preparing metadata (setup.py) ... done
Collecting pyzmq>=17.1.0 (from -r requirement.txt (line 5))
  Downloading pyzmq-25.1.1-cp39-cp39-macosx_10_15_universal2.whl.metadata (4.9 kB)
Collecting flask (from -r requirement.txt (line 6))
  Downloading flask-3.0.0-py3-none-any.whl.metadata (3.6 kB)
Collecting flask_compress (from -r requirement.txt (line 7))
  Downloading Flask_Compress-1.14-py3-none-any.whl.metadata (7.8 kB)
Collecting flask_json (from -r requirement.txt (line 8))
  Downloading Flask_JSON-0.4.0-py3-none-any.whl (8.7 kB)
Collecting Werkzeug>=3.0.0 (from flask->-r requirement.txt (line 6))
  Downloading werkzeug-3.0.1-py3-none-any.whl.metadata (4.1 kB)
Collecting Jinja2>=3.1.2 (from flask->-r requirement.txt (line 6))
  Downloading Jinja2-3.1.2-py3-none-any.whl (133 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 133.1/133.1 kB 13.3 MB/s eta 0:00:00
Collecting itsdangerous>=2.1.2 (from flask->-r requirement.txt (line 6))
  Downloading itsdangerous-2.1.2-py3-none-any.whl (15 kB)
Collecting click>=8.1.3 (from flask->-r requirement.txt (line 6))
  Downloading click-8.1.7-py3-none-any.whl.metadata (3.0 kB)
Collecting blinker>=1.6.2 (from flask->-r requirement.txt (line 6))
  Downloading blinker-1.7.0-py3-none-any.whl.metadata (1.9 kB)
Collecting importlib-metadata>=3.6.0 (from flask->-r requirement.txt (line 6))
  Downloading importlib_metadata-6.8.0-py3-none-any.whl.metadata (5.1 kB)
Collecting brotli (from flask_compress->-r requirement.txt (line 7))
  Downloading Brotli-1.1.0-cp39-cp39-macosx_10_9_x86_64.whl.metadata (5.5 kB)
Collecting zipp>=0.5 (from importlib-metadata>=3.6.0->flask->-r requirement.txt (line 6))
  Downloading zipp-3.17.0-py3-none-any.whl.metadata (3.7 kB)
Collecting MarkupSafe>=2.0 (from Jinja2>=3.1.2->flask->-r requirement.txt (line 6))
  Downloading MarkupSafe-2.1.3-cp39-cp39-macosx_10_9_x86_64.whl.metadata (3.0 kB)
Downloading pyzmq-25.1.1-cp39-cp39-macosx_10_15_universal2.whl (1.9 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.9/1.9 MB 56.0 MB/s eta 0:00:00
Downloading flask-3.0.0-py3-none-any.whl (99 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 99.7/99.7 kB 11.1 MB/s eta 0:00:00
Downloading Flask_Compress-1.14-py3-none-any.whl (8.4 kB)
Downloading blinker-1.7.0-py3-none-any.whl (13 kB)
Downloading click-8.1.7-py3-none-any.whl (97 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 97.9/97.9 kB 9.0 MB/s eta 0:00:00
Downloading importlib_metadata-6.8.0-py3-none-any.whl (22 kB)
Downloading werkzeug-3.0.1-py3-none-any.whl (226 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 226.7/226.7 kB 22.5 MB/s eta 0:00:00
Downloading Brotli-1.1.0-cp39-cp39-macosx_10_9_x86_64.whl (446 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 446.2/446.2 kB 37.9 MB/s eta 0:00:00
Downloading MarkupSafe-2.1.3-cp39-cp39-macosx_10_9_x86_64.whl (13 kB)
Downloading zipp-3.17.0-py3-none-any.whl (7.4 kB)
Building wheels for collected packages: GPUtil
  Building wheel for GPUtil (setup.py) ... done
  Created wheel for GPUtil: filename=GPUtil-1.4.0-py3-none-any.whl size=7392 sha256=1a36b3d366bc255e0dff9ae6efcd16047d977a360e99f94df
23b9496642f68e0
  Stored in directory: /Users/zhaokaifa_imac/Library/Caches/pip/wheels/2b/b5/24/fbb56595c286984f7315ee31821d6121e1b9828436021a88b3
Successfully built GPUtil
Installing collected packages: GPUtil, brotli, zipp, pyzmq, MarkupSafe, itsdangerous, click, blinker, Werkzeug, Jinja2, importlib-met
adata, flask, flask_json, flask_compress
```

- 4.2.2 Following the './BERT-BiLSTM-CRF-NER-tf-master/README.md' to download the pre-trained model and setup BERT-BiLSTM-CRF-NER. Besides, please make sure that the tensorflow (GPU version is strongly recommended) is correctly installed in your environment.
  - In the evaluation phase, we use the model (L=12, H=768, link: https://storage.googleapis.com/bert_models/2020_02_20/uncased_L-12_H-768_A-12.zip) for evaluation.
- 4.2.3 Please replace variable of path to pre-trained model in "run.sh" (line 5-7) to the path you downloaded the model in Step 4.2.2.

- run.sh

```
#export CUDA_VISIBLE_DEVICES=1,2,
bert-base-ner-train \
  -data_dir "./data/" \
  -output_dir "./results" \
  -init_checkpoint "./uncased_L-12_H-768_A-12/bert_model.ckpt" \
  -bert_config_file "./uncased_L-12_H-768_A-12/bert_config.json" \
  -vocab_file "./uncased_L-12_H-768_A-12/vocab.txt" \
  -num_train_epochs 100 \
  -do_train=true \
  -do_eval=true \
  -do_predict=true
```

- 4.2.4 Execute the following script in command line to reproduce the results:

```
sh run.sh
```

-

```
s 100 -do_train=true -do_eval=true -do_predict=true
                ARG   VALUE
_____________________________________________________
          batch_size = 64
    bert_config_file = ./uncased_L-12_H-768_A-12/bert_config.json
                cell = lstm
               clean = True
                clip = 0.5
            data_dir = ./data/
          device_map = 0
             do_eval = True
       do_lower_case = True
          do_predict = True
            do_train = True
        dropout_rate = 0.5
     filter_adam_var = False
     init_checkpoint = ./uncased_L-12_H-768_A-12/bert_model.ckpt
          label_list = None
       learning_rate = 1e-05
           lstm_size = 128
      max_seq_length = 128
                 ner = ner
          num_layers = 1
    num_train_epochs = 100.0
          output_dir = ./results
save_checkpoints_steps = 500
  save_summary_steps = 500
             verbose = False
          vocab_file = ./uncased_L-12_H-768_A-12/vocab.txt
   warmup_proportion = 0.1
```

-

```
processed 123043 tokens with 5484 phrases; found: 6541 phrases; correct: 2191.
accuracy:  82.46%; precision:  33.50%; recall:  39.95%; FB1:  36.44
                 : precision:   0.00%; recall:   0.00%; FB1:   0.00  11
          collect: precision:  23.21%; recall:   7.88%; FB1:  11.76  168
        condition: precision:   6.76%; recall:  24.93%; FB1:  10.63  2501
             data: precision:  52.01%; recall:  64.64%; FB1:  57.64  2386
          handler: precision:  67.62%; recall:  50.31%; FB1:  57.70  596
          purpose: precision:  38.57%; recall:  43.57%; FB1:  40.92  879
            share: precision:   0.00%; recall:   0.00%; FB1:   0.00  0
         subjects: precision:   0.00%; recall:   0.00%; FB1:   0.00  0
```

- 4.3 For LatticeLSTM:
  - 4.3.1 Please make sure the requirements in './Batch_Parallel_LatticeLSTM-master/requirement.txt', follow the '/Batch_Parallel_LatticeLSTM-master/README.md' to download the embeddings. Besides, please make sure that the pytorch (GPU version is strongly recommended) is correctly installed in your environment. Modify the paths in './Batch_Parallel_LatticeLSTM-master/pathes.py'.

```
cd ../
cd Batch_Parallel_LatticeLSTM-master/
conda create -n LaticeLSTM python=3.7.3
conda activate LaticeLSTM
pip install fastNLP
pip install pytorch>=1.1.0
pip install numpy>=1.16.4
pip install fitlog>=0.2.0
```

```
python main_CSP3.py
```

```
python main_CSP3.py
Auto commit by fitlog
device cuda
debug False
norm_embed False
batch 4
test_batch 512
optim sgd
lr 0.045
model lattice
skip_before_head False
hidden 113
momentum 0
bi True
dataset ppzh
use_bigram True
embed_dropout 0.5
gaz_dropout -1
output_dropout 0.5
epoch 100
seed 100
predict True
local_rank None
dict_keys(['train', 'dev', 'test'])
2059
1842
14678
Found 1791 out of 1814 words in the pre-training embedding.
Found 49252 out of 55271 words in the pre-training embedding.
Save cache to cache/zhppner.
Save cache to cache/load_yangjie_rich_pretrain_word_list.
```

```
 favor of nn.init.orthogonal_.
  nn.init.orthogonal(self.weight_ih.data)
/home/zachary/Code/EMNLP/CA4P-483/baselines/Batch_Parallel_LatticeLSTM-master/modules.py:44: UserWarning
avor of nn.init.constant_.
  nn.init.constant(self.bias.data, val=0)
bmeso
label_vocab:28
{0: 'O', 1: 'I-purpose', 2: 'I-data', 3: 'I-condition', 4: 'B-data', 5: 'E-data', 6: 'I-subjects', 7: 'B
, 10: 'B-condition', 11: 'E-handler', 12: 'E-condition', 13: 'I-handler', 14: 'B-collect', 15: 'B-share'
subjects', 19: 'E-subjects', 20: 'I-share', 21: 'I-collect', 22: 'S-handler', 23: 'S-condition', 24: 'S-
 27: 'S-collect'}
input fields after batch(if batch size is 2):
        chars: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
        target: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
        bigrams: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
        seq_len: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2])
        skips_l2r_source: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86, 2])
        skips_l2r_word: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86, 2])
        skips_r2l_source: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86, 2])
        skips_r2l_word: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86, 2])
        lexicon_count: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
        lexicon_count_back: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
target fields after batch(if batch size is 2):
        target: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2, 86])
        seq_len: (1)type:torch.Tensor (2)dtype:torch.int64, (3)shape:torch.Size([2])

training epochs started 2023-11-16-21-21-29-886966
Epoch 1/100:   0%|                                                                | 1735/367000 [27:51<114:
Epoch 1/100:   0%|                                                                             | 1770/367000
```

- 4.4 For the **Manual Agreements** given in Table3, please refer to the form: "./CA4P-483/Manual Kappa Evaluation 0611.xlsx".
  - In the form, row 1-9 denotes the manual check results of the 1st annotator, row 12-20 refers to the manual check results of the 2nd annotator, and row 56-64 gives the **Manual Agreements** based on the manual check results, i.e., row 1-20.
- *** Please note that the above algorithms may include random processes during operation, such as random initialization parameters, which may cause the results to be slightly different from those given in the paper.**

5. Reproduce Figure.4

Run the following script in command line to reproduce Figure.4:

```
python ./CA4P-483/DataSetStatistics/Figure4.py
```

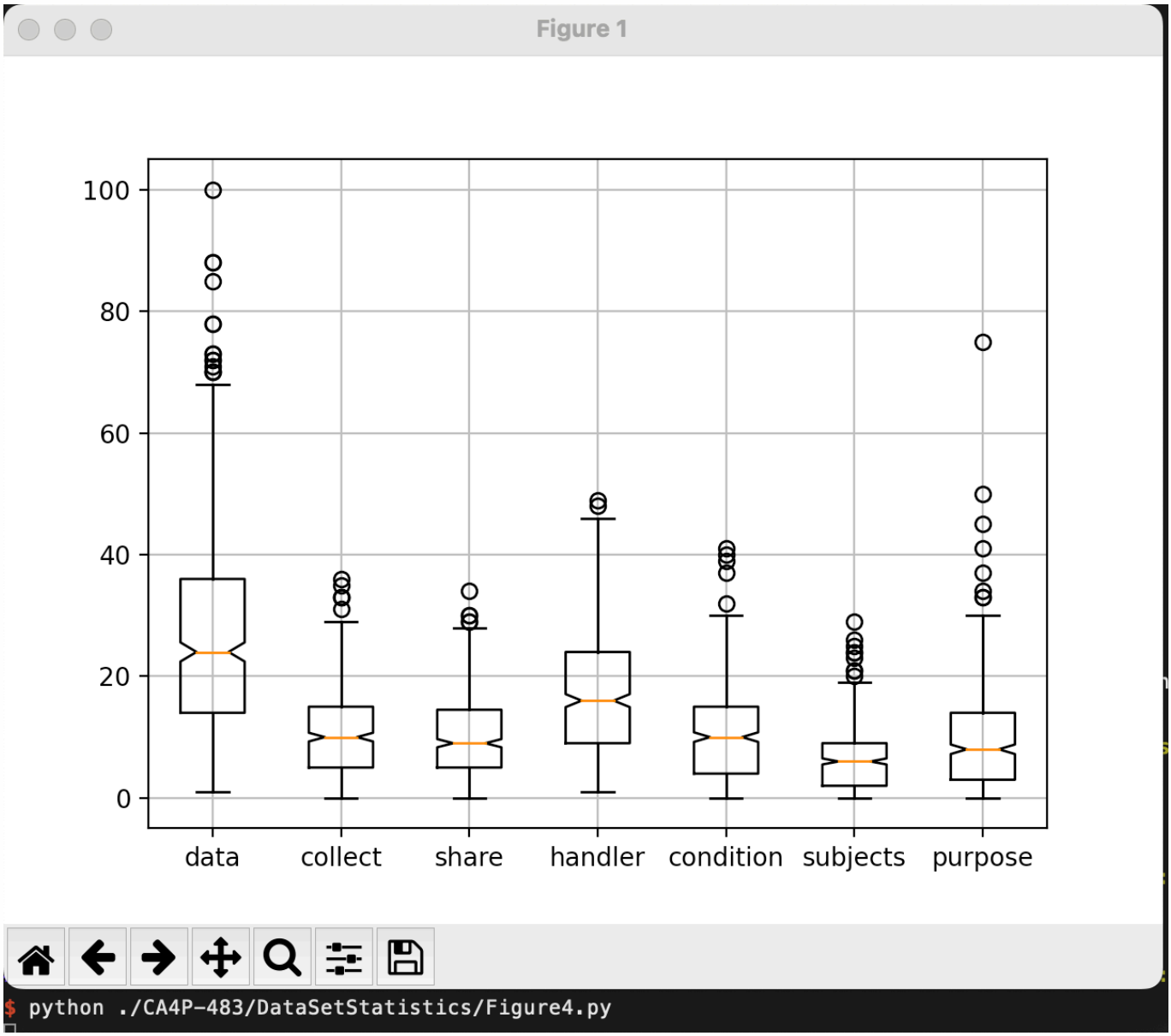
Figure 1
100
80
60
40
20
0
data
collect
share
handler
condition
subjects
purpose
$ python ./CA4P-483/DataSetStatistics/Figure4.py