Building Neural Net Machine Translation Systems Using Interlinear Glossed Texts

Yuan-Lu Chen

April 18, 2018

(1) a. Training Data:

Tha   a   athair nas   sine   na   a   mhàthair.
be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother

'His father is older than his mother.'

b. Test Data:

Tha   mi a' sireadh   an   leabhair bhig
be-PRES-IND 1S PROG searching-VN ART book-G small-G
ghuirm.
blue-G

'I am looking for the small blue book.'

(2) **Gloss-helps hypothesis: the translation systems trained with the gloss data incorporated should outperform the systems trained with only Gaelic and English sentences pairs (i.e. without gloss data).**
The hypothesis can have two versions, strong and weak:

a. Strong version: Gloss may replace the source natural language totally, and the system outperforms the system trained with source natural language to target language sentence pairs (i.e. the baseline systems).

b. Weak version: Gloss only increases the performance of the baseline systems, but cannot replace the source language.

(3) Baseline:

a. Trained with:
Gaelic sentence $\rightarrow$ English Sentence

   i. e.g.
   Tha a athair nas sine na a mhàthair.
   $\rightarrow$
   His father is older than his mother.

b. Test with:
Tha mi a' sireadh an leabhair bhig ghuirm.
$\rightarrow$
Predicted_English_Translation$_j$

    c. Evaluation of the model:
       compare
       Predicted_English_Translation$_j$
       with
       'I am looking for the small blue book.'

(4) Gloss Treatment:
    a. Trained with:
       Gloss line → English Sentence

       i. e.g.
          be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother
          →
          His father is older than his mother.

    b. Test with:
       be-PRES-IND 1S PROG searching-VN ART book-G small-G blue-G
       →
       Predicted_English_Translation$_j$

Round	Gaelic (Baseline)	GLOSS
0	17.29	18.39
1	16.42	18.00
2	15.29	16.02
3	15.97	20.22
4	17.79	19.02
5	16.73	15.53
6	17.11	18.00
7	16.37	20.08
8	15.93	15.82
9	16.99	15.93
Mean	16.59	17.70

Table 1: BLEU scores of Model$_{GDtoEN}$ and Model$_{GLOSStoEn}$

The average score of the Models$_{GLOSStoEN}$ is only slightly higher than the average score of the Models$_{GDtoEN}$. Also, after doing a paired T-test, the difference between the two types of models is not attested (M_{GDToEn}=16.59, SD_{GDToEn}=0.74; $M_{GLOSStoEN}$=17.70, $SD_{GLOSStoEN}$=1.78; $t(9)$=1.97, p=0.080)

(5) Parallel-Partial Treatment:
    a. Trained with:
       Gaelic sentence → English Sentence
       Gloss line → English Sentence
       Gloss line → Gaelic sentence
       Gaelic word → Gloss

    b. Parallel
       i. Gaelic to English:
          <"Tha a athair nas sine na a mhàthair", "His father is older than his mother.">

   ii. Gloss to English:
      <"be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother",
      "His father is older than his mother">

   iii. Gloss to Gaelic:
      <"be.pres 3sm.poss father comp old.cmpr comp 3sm.poss mother",
      "Tha a athair nas sine na a mhàthair">

c. Partial

   i. <"Tha", "be.pres">

   ii. <"a", "3sm.poss">

   iii. <"athair", "father">

   iv. <"nas", "comp">

   v. <"sine", "old.cmpr">

   vi. <"na", "comp">

   vii. <"a", "3sm.poss">

   viii. <"mhàthair", "mother">

d. Test with:
Tha mi a' sireadh an leabhair bhig ghuirm.
$\rightarrow$
Predicted_English_Translation$_j$

Round	Gaelic (Baseline)	ParaPart
0	17.29	32.64
1	16.42	32.28
2	15.29	29.94
3	15.97	31.18
4	17.79	32.83
5	16.73	31.11
6	17.11	32.19
7	16.37	33.52
8	15.93	30.93
9	16.99	34.35
Mean	16.59	32.10

Table 2: BLEU scores of Model$_{\text{GDtoEN}}$ and Model$_{\text{ParaParttoEn}}$

$M_{\text{GDToEn}}=16.59$, $SD_{\text{GDToEn}}=0.74$; $M_{\text{ParaPart}}=32.10$, $SD_{\text{ParaPart}}=1.33$; $t(9)=48.95$, $p<0.01$.

(6) Many other possible ways to blend Gaelic and gloss:

Round	Baseline	GLOSS	ParaPart	Para	Interleaving	Concat	Google Translation
0	17.29	18.39	32.64	25.42	13.67	15.42	22.09
1	16.42	18.00	32.28	25.32	12.49	14.31	25.38
2	15.29	16.02	29.94	20.72	11.01	15.38	23.72
3	15.97	20.22	31.18	22.22	12.33	14.18	23.21
4	17.79	19.02	32.83	24.27	12.56	18.63	22.31
5	16.73	15.53	31.11	24.55	12.13	14.89	23.41
6	17.11	18.00	32.19	27.03	11.55	15.16	24.53
7	16.37	20.08	33.52	25.34	12.78	15.20	22.78
8	15.93	15.82	30.93	24.24	12.43	15.50	25.67
9	16.99	15.93	34.35	25.96	11.65	15.72	23.42
Mean	16.59	17.70	32.10	24.51	12.26	15.44	23.65

Table 3: BLEU scores of the treatments: Ten rounds of repeated random sub-sampling validation are conducted. For each round, the same sets of IGTs are used. Each column is a treatment, and each row is a single round of repeated random sub-sampling validation. The last column is the scores of Google Translation. We used a free Google translation API (Han, 2018) to translate the same set of test Gaelic sentences into English.

References

Han, SuHun (2018), "Googletrans." https://github.com/ssut/py-googletrans.