HOO 2012: Helping Our Own Shared Task

Vidas Daudaravičius
Vytautas Magnus University
v.daudaravicius@if.vdu.lt

21th June 2012

Outline

HOO task

- Proposed to address automated error correction of non-native English texts
- For the scientific papers in the computational linguistics domain.

HOO 2012 task

- Helping Our Own 2012 Shared Task focuses on
  determiners and **prepositions** as they are well-known
  sources for errors produced by non-native English writers.
- To provide a forum for the comparative evaluation of
  different approaches to the correction of these errors
- Hosted at the 7th Building Educational Applications
  Workshop, NAACL, Montreal, Jun 7.

HOO 2012 error correction types

- Replace determiner (RD):
  *Have **the** nice day.* $\mapsto$ *Have **a** nice day.*

- Replace preposition (RT):
  ***In** the other hand...* $\mapsto$ ***On** the other hand...*

- Missing determiner (MD):
  That is great idea. $\mapsto$ *That is **a** great idea.*

- Missing preposition (MT):
  She woke up 6 oclock. $\mapsto$ *She woke up **at** 6 oclock.*

- Unnecessary determiner (UD):
  *I like **the** pop music.* $\mapsto$ *I like pop music.*

- Unnecessary preposition (UT):
  *He must go **to** home.* $\mapsto$ *He must go home.*

Outline

HOO 2012 Data

- Training data:
  - Subset of the CLC FCE Dataset (1244 scripts).
  - 1000 exam scripts
  - Written by students who attended the Cambridge ESOL First Certificate in English examination.
  - 374,680 words, with an average of 375 words per text.
- Test data:
  - 100 new previously unseen texts
  - 18013 words.

HOO 2012 Data - Raw text

< PART id="1" >
Dear Miss Helen Ryan,
I am writing to you because I am very glad for my first place in the competition,
I am going to tell you the things that you ask me on your letter.
I would like to travel on July because is the perfect time to go to the camp and
is summer.
Accommodation at camp I would like to choose cabins because I think I would
be protected from the animals and the cold.
Furthermore for the activities I want to select Tennis and Basketball because I
haves been playing tennis since I was young and basketball because I played
for the team in my college like the capitan.
I want to ask you one more thing I do not have any idea about how much
money and clothes I have to bring would you give me an advice to know how
much I need to bring
I look forward to heard from you
Yours sincerely
< /PART >

HOO 2012 Data - Standoff

```
<edit end="67" file="0082" index="0001" part="1" start="64" type="RT">
 <original>for</original>
 <corrections>
  <correction>about</correction>
 </corrections>
</edit>
<edit end="126" file="0082" index="0002" part="1" start="126" type="MT">
 <original>
  <empty/>
 </original>
 <corrections>
  <correction>about </correction>
 </corrections>
</edit>
```

HOO 2012 Data

Type	# Training	# Test
UT	822	43
MT	1104	57
RT	2618	136
Prep	4545	236
UD	1048	53
MD	2230	125
RD	609	39
Det	3887	217
Prep + Det Total	8432	453
Words per Error	44.18	39.77

Outline

Evaluation Metrics

Detection : does the system determine that an edit of the specified type is required at some point in the text?

Recognition : does the system correctly determine the extent of the source text that requires editing?

Correction : does the system offer a correction that is identical to that provided in the gold standard?

Evaluation Criteria

Precision

$$precision = \frac{tp}{tp + fp}$$

Recall

$$recall = \frac{tp}{tp + fn}$$

F-measure

$$F = \frac{2 * precision * recall}{precition + recall}$$

Outline

Participating Teams

Team ID	Group or Institution	Runs
CU	Computer Laboratory, University of Cambridge, **UK**	8
ET	Educational Testing Service, New Jersey, **USA**	3
JU	Jadavpur University, Kolkata, **India**	1
KU	Natural Language Processing Lab, Korea University, Seoul, **Korea**	10
LE	KU Leuven, **Belgium**	2
NA	NAIST, **Japan**	8
NU	National University of Singapore, **Singapore**	1
TC	Department of Computer Science and Statistics, Trinity College Dublin, **Ireland**	10
TH	NLPLAB, National Tsing Hua University, Hsinchu, **Taiwan**	4
UD	UKP Lab, Technische Universitat Darmstadt, **Germany**	3
UI	Cognitive Computation Group, University of Illinois, **USA**	10
UT	Theoretical Computational Linguistics Group, University of Tubingen, **Germany**	10
VA	Valkuil.net, **The Netherlands**	6
VT	VTEX, Vilnius, **Lithuania**	9

Outline

Word length and word ending transformations

word – keeps the whole word (e.g. *make* $\mapsto$ *make*);

2end – takes the length of a word and adds the last two characters (*make* $\mapsto$ *4_ke*);

1end – takes the length of a word and adds the last character (*make* $\mapsto$ *4_e*).

Error correction rules

L_2	L_1	type	original	correction	R_1	R_2
	asked	RT	for	about	the	
i	asked	RT	for	about		
to	asked	RT	for	about		
		RT	for	about	the	camp
		RT	for	about	the	discounts
		RT	for	about	the	experience
		RT	for	about	the	first
		RT	for	about	the	new
		RT	for	about	the	news
		RT	for	about	the	play
		RT	for	about	the	prise
		RT	for	about	the	terrible
		RT	for	about	the	very

Ranking Possible Corrections

- Knesser–Ney trigram smoothing
- sentence probability $p(x) = \prod_i \hat{p}(x_i \mid x_{i-2}, x_{i-1})$
- OC ratio:

$$OC\ ratio(x) = \frac{p(x|O)}{p(x|C)}$$

where $p(x|O)$ is the probability of the sentence x using the original language model
$p(x|C)$ is the probability of the sentence x using the corrected language model

Ranking Possible Corrections

Doc ID	Run	Rules applied	OC ratio	sentence correction
2025	2	the//MD/ that/this/RD/	0.451	is the 8 th july till **the** end of **that** month , what do you think ?
		that/this/RD/	0.559	is the 8 th july till end of **that** month , what do you think ?
		the//MD/	0.633	is the 8 th july till **the** end of this month , what do you think ?
		–	0.785	is the 8 th july till end of this month , what do you think ?
	7	the//MD/ that/this/RD/	0.345	3s the 8 2h 4y till **the** 3d of **that** 5h , what 2o you 5k ?
		that/this/RD/	0.441	3s the 8 2h 4y till 3d of **that** 5h , what 2o you 5k ?
		the//MD/	0.533	3s the 8 2h 4y till **the** 3d of this 5h , what 2o you 5k ?
		–	0.683	3s the 8 2h 4y till 3d of this 5h , what 2o you 5k ?
2043	2	/for/UT/	0.976	i am writing in response to your last letter , to answer and ask you ~~for~~ some questions .
		–	1.035	i am writing in response to your last letter , to answer and ask you for some questions .
	7	/for/UT/	0.966	i am 7g in 8e to your 4t 6r , to 6r and 3k you ~~for~~ some 9s .
		a//MD/ /for/UT/	1.022	i am 7g in **a** 8e to your 4t 6r , to 6r and 3k you ~~for~~ some 9s .
		–	1.025	i am 7g in 8e to your 4t 6r , to 6r and 3k you for some 9s .
		a//MD/	1.085	i am 7g in **a** 8e to your 4t 6r , to 6r and 3k you for some 9s .

Outline

All Corrections

Team	Run	Precision	Recall	F-Score
CU	0	**70.00**	4.64	8.70
ET	0	0.00	0.00	0.00
JU	1	2.52	2.65	2.58
KU	0	2.11	22.52	3.86
LE	0	31.15	22.08	25.84
NA	3	29.43	20.53	24.19
NU	0	45.45	20.97	**28.70**
TC	9	2.66	12.80	4.41
TH	1	9.44	25.61	13.79
UD	2	1.20	4.19	1.87
UI	1	26.39	**28.26**	27.29
UT	7	21.95	15.89	18.44
VA	3	6.16	7.51	6.77
VT	5	8.76	4.19	5.67

Determiner Corrections

Team	Run	Precision	Recall	F-Score
CU	6	5.25	33.18	9.07
ET	0	0.00	0.00	0.00
JU	1	1.66	1.38	1.51
KU	0	6.15	5.07	5.56
LE	0	34.72	30.88	32.68
NA	1	28.88	24.88	26.73
NU	0	**48.28**	25.81	33.63
TC	3	4.11	4.15	4.13
TH	1	9.91	19.35	13.10
UD	0	0.00	0.00	0.00
UI	0	35.61	**33.64**	**34.60**
UT	6	25.75	19.82	22.40
VA	3	5.02	6.91	5.81
VT	5	7.98	5.99	6.84

Preposition Corrections

Team	Run	Precision	Recall	F-Score
CU	0	**61.11**	4.66	8.66
ET	0	0.00	0.00	0.00
JU	1	3.05	3.81	3.39
KU	0	1.27	25.00	2.41
LE	0	25.78	13.98	18.13
NA	3	30.23	16.53	21.37
NU	0	41.94	16.53	**23.71**
TC	9	3.06	13.98	5.02
TH	1	9.19	**31.36**	14.22
UD	1	7.63	4.24	5.45
UI	1	20.36	24.15	22.09
UT	0	25.45	11.86	16.18
VA	0	7.51	8.05	7.77
VT	7	12.24	2.54	4.21

Error Correction for Medical Domain

- Patient Health record language
- Correct term usage
- Correct Named Entity usage (address, time ...)
- Creating and Spreading best-practices of medical language usage.

HOO 2012: Helping Our Own Shared Task

Vidas Daudaravičius
Vytautas Magnus University
v.daudaravicius@if.vdu.lt

21th June 2012