

Easier Automatic Sentence Simplification Evaluation

Fernando Alva-Manchego, Louis Martin, Carolina Scarton and Lucia Specia

WHAT IS SENTENCE SIMPLIFICATION?

To modify the content and structure of a sentence so that it is **easier to understand** while keeping most of its original meaning

Original

The second **largest** city of Russia ~~and one of the world's major cities~~, St . Petersburg has played a **vital** role in Russian history.

Simplification

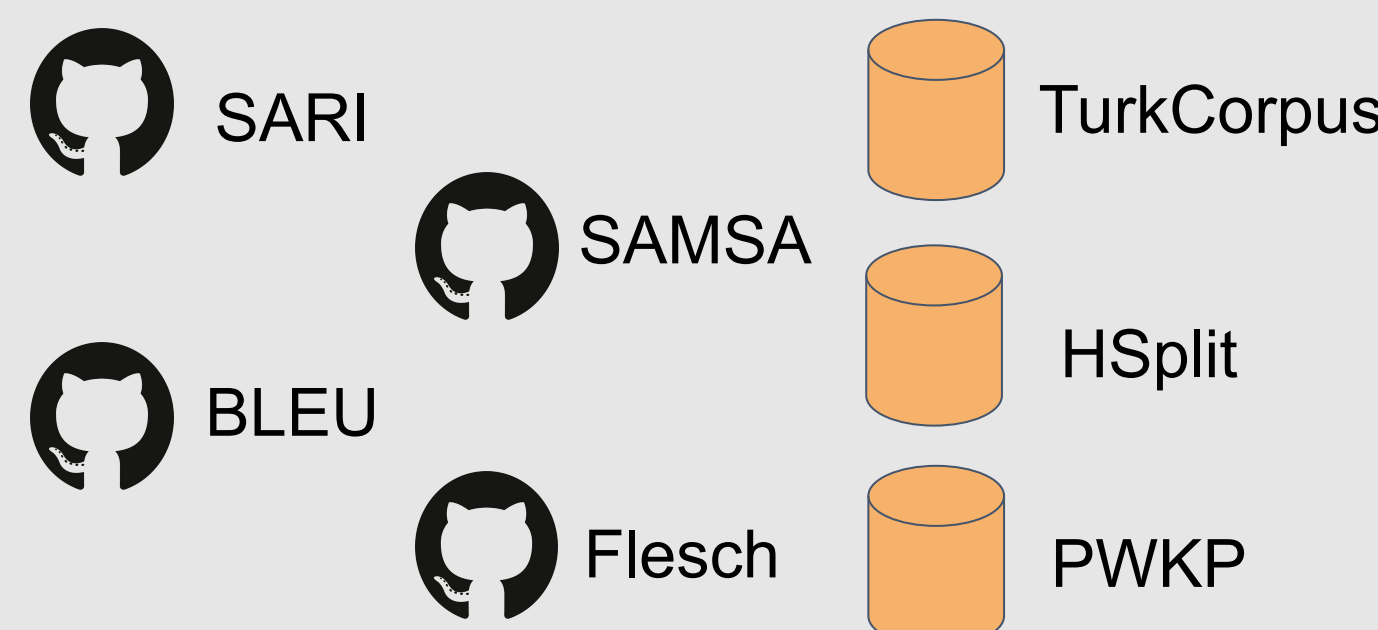
St. Petersburg is the second **biggest** city of Russia.
St. Petersburg has played an **important** role in Russian history.



- Uncommon words are replaced
- Complex syntactic structures are changed (e.g. sentence splitting)
- “Unimportant” information is removed

PROBLEMS WITH AUTOMATIC EVALUATION

Metrics and datasets are scattered in different repositories



...and they are not enough for comparing systems

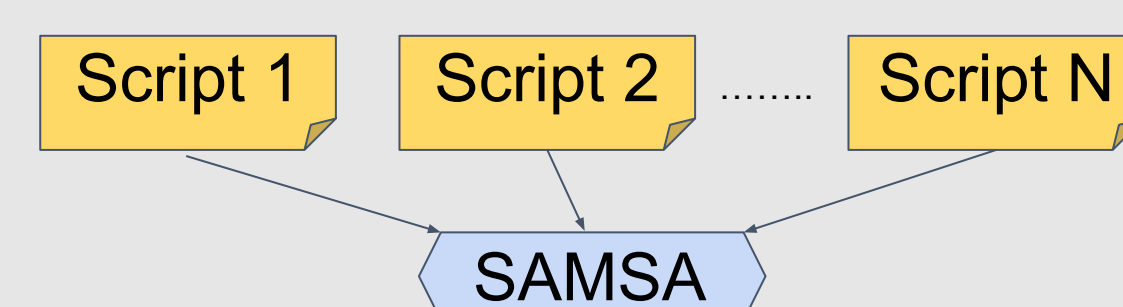
how can I compare my model with previous models with standard metrics?

Which text transformations are performed by the model?

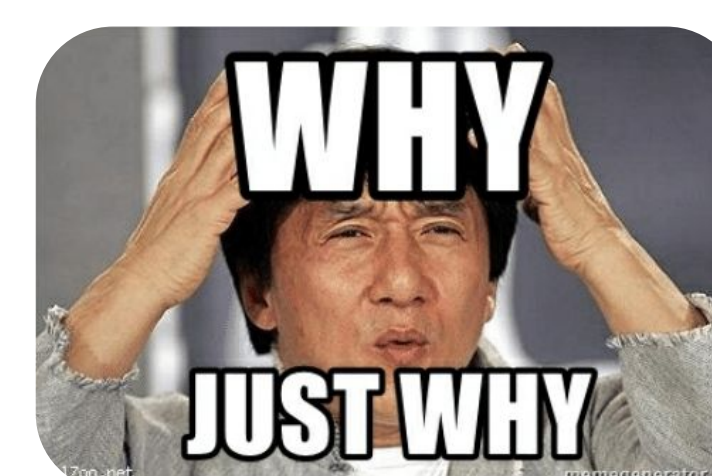
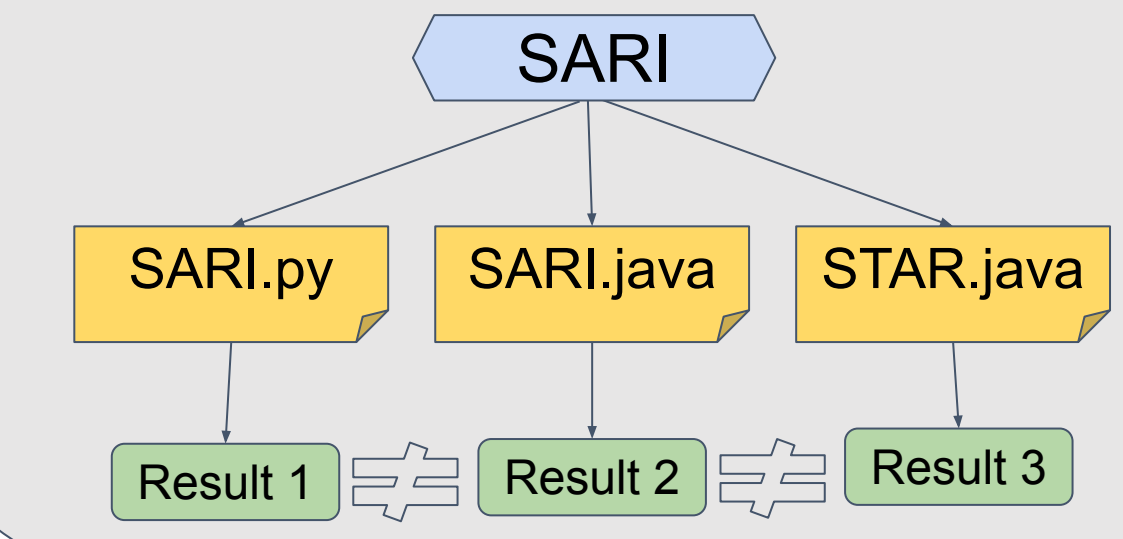


How can I easily analyse my system's output?

Some metrics are convoluted to run...



...others have multiple incompatible implementations



<https://github.com/feralvam/easse>



```
(dev) ~$ easse evaluate -t turk -m 'bleu,sari' -q < DMASS-DCSS.tok.low
BLEU: 73.29
SARI: 40.42
Quality estimation: {'Compression ratio': 0.89, 'Sentence splits': 0.98, 'Levenshtein similarity': 0.82, 'Exact matches': 0.05, 'Additions proportion': 0.15, 'Deletions proportion': 0.21, 'Lexical complexity score': 8.06}
(dev) ~$
```

```
>>> from easse.sari import corpus_sari
>>> orig_sents = ["About 95 species are currently accepted ."]
>>> sys_sents = ["About 95 you now get in ."]
>>> refs_sents = [{"About 95 species are currently known .",
...               "About 95 species are now accepted .",
...               "95 species are now accepted ."}]
>>> sari_score = corpus_sari(orig_sents, sys_sents, refs_sents)
>>> print(sari_score)
27.381587381587384
```



Standard Metrics

System	TurkCorpus		HSplit
	SARI	BLEU	SAMS
Reference	49.88	97.41	54.00
PBSMT-R	38.56	81.11	47.59
Hybrid	31.40	48.97	46.68
SBSMT-SARI	39.96	73.08	41.41
DRESS-Ls	37.27	80.12	45.94
DMASS-DCSS	40.42	73.29	35.45

Reference-less Quality Estimation Features

System	Compression ratio	Exact matches	Additions proportion	Deletion proportion
PBSMT-R	0.95	0.1	0.1	0.11
Hybrid	0.57	0.03	0.01	0.41
SBSMT-SARI	0.94	0.11	0.16	0.13
DRESS-Ls	0.78	0.26	0.04	0.26
DMASS-DCSS	0.89	0.05	0.15	0.21

Performance Analysis Report

EASSE report

Scores

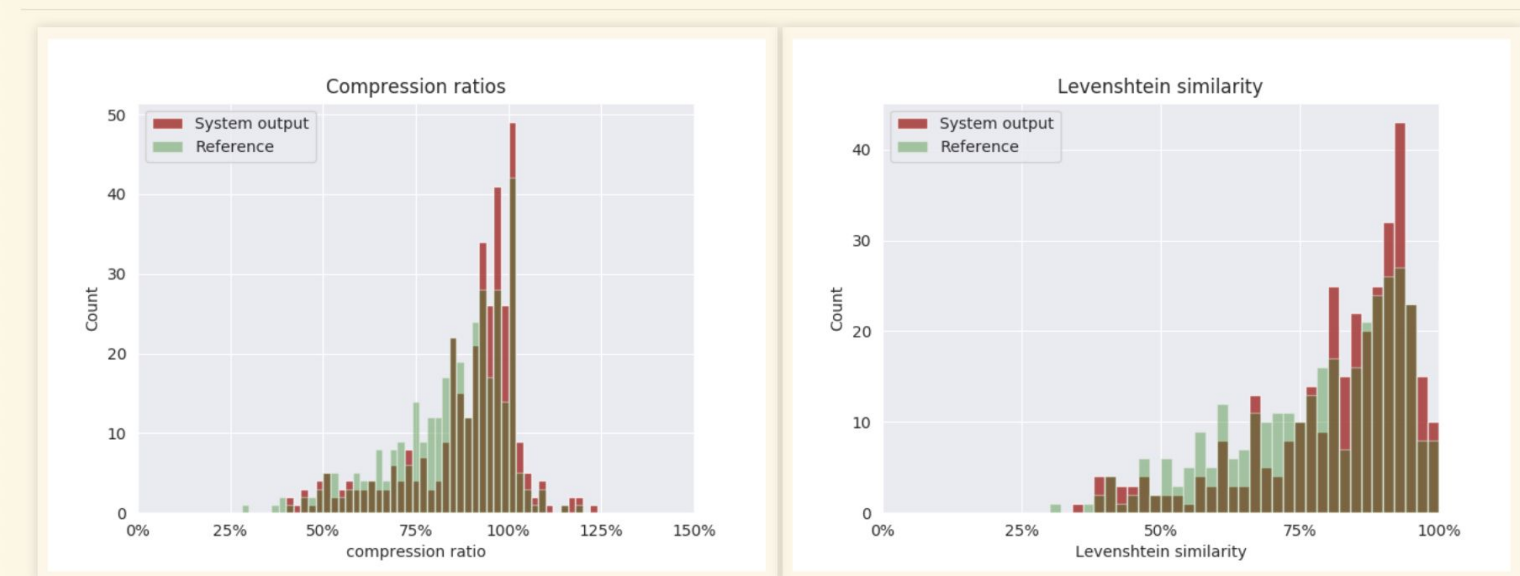
System vs. Reference

	BLEU	SARI	FKGL	Compression ratio	Sentence splits	Levenshtein similarity	Exact matches	Additions proportion	Deletions proportion	Lexical complexity score
System output	73.29	40.42	7.66	0.89	0.98	0.82	0.05	0.15	0.21	8.44
Reference	68.02	41.31	8.22	0.85	1.01	0.8	0.08	0.13	0.26	8.59

By sentence length (characters)

	BLEU	SARI	FKGL	Compression ratio	Sentence splits	Levenshtein similarity	Exact matches	Additions proportion	Deletions proportion	Lexical complexity score
length: [25,71]	68.10	41.76	3.74	0.87	1.00	0.80	0.08	0.19	0.25	8.28
length: [71,102]	70.04	40.45	5.58	0.89	1.02	0.82	0.07	0.15	0.21	8.44
length: [102,191]	74.61	39.38	7.55	0.92	0.95	0.85	0.06	0.15	0.18	8.28
length: [191,147]	70.07	40.85	8.89	0.88	0.96	0.81	0.04	0.16	0.24	8.49
length: [147,344]	77.48	40.01	11.06	0.89	0.95	0.85	0.01	0.10	0.18	8.71

Plots



Randomly sampled simplifications

one side of the armed conflicts is composed mainly of the sudanese military and the jarjaweed , a sudanese militia group recruited mostly from the afro-arab abakka tribe of the northern coast region in sudan .
one side of the armed conflicts is composed mainly of the sudanese military and the jarjaweed , a sudanese militia group recruited mostly from the afro-arab tribes of the northern coast region in sudan .
jeddah is the petrodollar gateway to mecca , islam 's holiest city , which **able-bodied** muslims are required to visit at least once in their lifetime .
jeddah is the main gateway to mecca , islam 's holiest city , which **physically** muslims are required to visit at least once in their lifetime .
the great dark spot is thought to **represent** a hole in the methane cloud deck of neptune .
the great dark spot is thought to **have** a hole in the methane cloud deck of neptune .
his next work , **saturday** , **follows** an especially eventful day in the life of a successful neurosurgeon .
his next work , **saturday** , **shows** an especially eventful day in the life of a successful neurosurgeon .
the tarantula , the trickster character , spun a black cord and , attaching it to the ball , **crawled** away fast to the east , **pulling** on the cord with all his strength .
the tarantula , the trickster character , spun a black cord and , attaching it to the ball , **crawled** away fast to the east , **pulling** on the cord with all his strength .
there he died six weeks later , on 13 january 888 .
he died six weeks later , on 13 january 888 .
they are culturally akin to the coastal **peoples** of papua new guinea .
they are culturally akin to the coastal **communities** of papua new guinea .
since 2000 , the recipient of the kate greenaway medal has also been presented with the colin means award to the value of £ 5000 .
since 2000 , the recipient of the kate greenaway medal has also been presented with the colin means award to the value of £ 5000 .
following the drummers are dancers , who often play the sogo (a tiny drum that makes almost no sound **-rds-** and tend to have more elaborate — even acrobatic — **choreography** .
following the drummers are dancers , who often play the sogo (a tiny drum that makes almost no sound **-rds-** and tend to have more elaborate — even acrobatic — **dances** .
the spacecraft consists of two main elements : the **mass cassini orbiter** , named after the italian-french astronomer giovanni domenico cassini , and the **esa huygens probe** , named after the dutch astronomer , mathematician and physicist christian huygens .
the spacecraft is named after the italian-french astronomer giovanni huygens .
Simplifications with only one differing word
the great dark spot is thought to **represent** a hole in the methane cloud deck of neptune .
the great dark spot is thought to **have** a hole in the methane cloud deck of neptune .
his next work , **saturday** , **follows** an especially eventful day in the life of a successful neurosurgeon .
his next work , **saturday** , **shows** an especially eventful day in the life of a successful neurosurgeon .
they are culturally akin to the coastal **peoples** of papua new guinea .
they are culturally akin to the coastal **communities** of papua new guinea .
it was **originally** thought that the debris thrown up by the collision filled in the smaller craters .
it was **first** thought that the debris thrown up by the collision filled in the smaller craters .
many species had **vanished** by the end of the nineteenth century , with european settlement .
many species had **gone** by the end of the nineteenth century , with european settlement .
in 1987 weiler was inducted into the rock and roll hall of fame .
in 1987 , weiler was inducted into the rock and roll hall of fame .
in its pure form , dextrometorphan **occurs** as a white powder .
in its pure form , dextrometorphan **happens** as a white powder .
admission to langhua is **extremely** competitive .
admission to langhua is **very** competitive .
for example , king bhutnabi was born on monday , so on his birthday throughout thailand will be decorated with yellow color .
for example , king bhutnabi was born on monday , so on his birthday throughout thailand , will be decorated with yellow color .

Word-level Transformations Accuracy

Hershey **left** no **heirs** when he died in 1945 , **giving** to charity **most** of his fortune .
Hershey died in 1945 and **gave** **most** of his fortune to charity .

System	Delete	Move	Replace	Copy
PBSMT-R	34.18	2.64	23.65	93.50
Hybrid	49.46	7.37	1.03	70.73
SBSMT-SARI	28.42	1.26	37.21	92.89
DRESS-Ls	40.31	1.43	12.62	86.76
DMASS-DCSS	38.03	5.10	34.79	86.70