NeuralREG: an end-to-end approach for Referring Expression Generation

Thiago Castro Ferreira1
Diego Moussallem2
Ákos Kádár1
Emiel Krahmer1
Sander Wubben1

TiCC - Tilburg University1
AKSW Research Group, University of Leipzig, Germany2

Supported by the National Council of Scientific and Technological Development from Brazil (CNPq).
The Aarhus Airport is located in Aarhus, Denmark. It is situated 25.0 meters above sea level. The airport has a runway called 10R/28L.
The Aarhus Airport is located in Aarhus, Denmark. It is situated 25.03 meters above sea level. The airport has a runway called 10R/28L.
MOTIVATION

Novel "end-to-end" NLG models
Generation of delexicalized templates from different meaning representations...

AMR → template → text
(Konstas et al., 2017)
(Castro Ferreira et al., 2017)

Dialog Act → template → dialogue text
(Wen et al., 2015)
(Dušek and Jurčíček, 2016)

RDF triples → template → text
WebNLG Challenge (Gardent et al., 2017)

...for accounting data sparsity and unseen entities
(Konstas et al., 2017)
DATA

WebNLG corpus
25,298 text describing 9,674 triple sets
Manually delexicalized
SUBJECT-1 is located in OBJECT-1. SUBJECT-1 is situated OBJECT-2 meters above sea level. SUBJECT-1 has a runway called OBJECT-3.
Aarhus_Airport is located in Aarhus,_Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L.

Conversion in constant time
**GOAL**

Aarhus_Airport is located in Aarhus, Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L.

↓REG

The Aarhus Airport is located in Aarhus, Denmark. It is situated 25.0 meters above sea level. The airport has a runway called 10R/28L.

Underestimated process so far.
**PROBLEM**

*Aarhus Airport* is located in Aarhus, Denmark. *Aarhus Airport* is situated 25.0 meters above sea level. *Aarhus Airport* has a runway called 10R/28L.

**vs.**

*The Aarhus Airport* is located in Aarhus, Denmark. *It* is situated 25.0 meters above sea level. *The airport* has a runway called 10R/28L.

REG is crucial for the coherence of the text.
REG MODELS

Extensively studied in pipeline architectures of NLG
GREC Challenges (Belz et al., 2010)

Decisions taken by different subtasks (modular)
Choice of referential form
Surface realization

Bottlenecks
Feature engineering
Difficulties in developing and maintaining
Propagation of errors in cascade along the modules
INPUT

**Target**
Target reference to be realized

**Pre-context**
Lowercased, tokenized and delexicalized piece of text *before* the target reference

**Pos-context**
Lowercased, tokenized and delexicalized piece of text *after* the target reference
Aarhus Airport is located in Aarhus, Denmark. Aarhus Airport is situated 25.0 meters above sea level. Aarhus Airport has a runway called 10R/28L.
EOS Aarhus_Airport is located in Aarhus, Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L. EOS
Aarhus Airport is located in Aarhus, Denmark. Aarhus Airport is situated 25.0 meters above sea level. Aarhus Airport has a runway called 10R/28L. It
EOS Aarhus_Airport is located in Aarhus, Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L. EOS
EOS Aarhus_Airport is located in Aarhus,_Denmark . Aarhus_Airport is situated 25.0 meters above sea level . Aarhus_Airport has a runway called 10R/28L . EOS

The airport
EOS Aarhus_Airport is located in Aarhus, Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L. EOS

Pre-context
Target
Pre-context

↓

10R/28L
NEURALREG

Encoder Attention-Decoder architecture

Context encoders
Vector representations for pre- and pos-contexts

Decoder
Combining representations and decoding the referring expression
NEURALREG

EOS Aarhus_Airport is located in Aarhus, Denmark.
Pre-context

Aarhus_Airport
TARGET

is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L. EOS
Pos-Context
Pre-context

EOS Aarhus_Airport is located in Aarhus, Denmark.

Pos-context

is situated 25.0 meters above sea level. ....
EOS Aarhus_Airport is located in Aarhus, Denmark.

is situated 25.0 meters above sea level.
EOS Aarhus Airport is located in Aarhus, Denmark.

is situated 25.0 meters above sea level.
EOS Aarhus Airport is located in Aarhus, Denmark.

is situated 25.0 meters above sea level.
Pre-context

EOS Aarhus Airport is located in Aarhus, Denmark.

Pos-context

is situated 25.0 meters above sea level.
Pre-context

EOS Aarhus Airport is located in Aarhus, Denmark.

Pos-context

is situated 25.0 meters above sea level.
Pre-context

*EOS Aarhus_Airport* is located in *Aarhus, Denmark*.

Pos-context

is situated **25.0** meters above sea level. *Aarhus_Airport* has a runway called **10R/28L**. *EOS*
EOS Aarhus_Airport is located in Aarhus, Denmark.

Aarhus_Airport is situated 25.0 meters above sea level.

Aarhus_Airport has a runway called 10R/28L.
EOS Aarhus_Airport is located in Aarhus, Denmark.

Aarhus_Airport is situated 25.0 meters above sea level.

Aarhus_Airport has a runway called 10R/28L.
EOS Aarhus_Airport is located in Aarhus, Denmark.
Aarhus_Airport has a runway called 10R/28L.
EOS is situated 25.0 meters above sea level.
DECODER

\[ s_i = \Phi_{\text{dec}}(s_{i-1}, [c_i, V_{y_{i-1}}, V_{\text{target}}]) \]

\[ y_i = \text{beam}(\text{softmax}(W_c s_i + b)) \]

evaluation of 3 methods to compute \( c_i \)...
SEQ2SEQ

Average and concat matrixes $h^{(pre)}$ and $h^{(pos)}$

$$\hat{h}^{(k)} = \frac{1}{N} \sum_{i}^{N} h^{(k)}_{i}$$

$$c_{i} = [\hat{h}^{(pre)}, \hat{h}^{(pos)}]$$
CATT

Concatenative attention

\[ e_{ij}^{(k)} = v_{a}^{(k)T} \tanh(W_{a}^{(k)} s_{i-1} + U_{a}^{(k)} h_{j}^{(k)}) \]

\[ \alpha_{ij}^{(k)} = \frac{\exp(e_{ij}^{(k)})}{\sum_{n=1}^{N} \exp(e_{in}^{(k)})} \]

\[ c_{i}^{(k)} = \sum_{j=1}^{N} \alpha_{ij}^{(k)} h_{j}^{(k)} \]

\[ c_{i} = [c_{i}^{(pre)}, c_{i}^{(pos)}] \]
Hierarchical Attention
(Libovický and Helcl, 2017)

\[ e_i^{(k)} = v_b^{(k)T} \tanh(W_b^{(k)} s_{i-1} + U_b^{(k)} c_i^{(k)}) \]

\[ \beta_i^{(k)} = \frac{\exp(e_i^{(k)})}{\sum_n \exp(e_i^{(n)})} \]

\[ c_i = \sum_k \beta_i^{(k)} U_b^{(k)} c_i^{(k)} \]
NEURALREG

\[ s_i = \Phi_{\text{dec}}(s_{i-1}, [c_i, V_{y_{i-1}}, V_{\text{target}}]) \]

**NeuralREG+Seq2Seq**

\[ c_i = [\text{avg}(h^{(pre)}), \text{avg}(h^{(pos)})] \]

**NeuralREG+CAtt**

\[ c_i = [\text{attend}(h^{(pre)}), \text{attend}(h^{(pos)})] \]

**NeuralREG+HierAtt**

\[ c_i = \text{hierattend}(\text{attend}(h^{(pre)}), \text{attend}(h^{(pos)})) \]
EVALUATION

WebNLG corpus
25,298 text describing 9,674 triple sets
Manually delexicalized

78,901 references to 1,483 entities
Train: 63,031 - Dev: 7,127 - Test: 8,743
BASELINES

Only Names

Ferreira
Aarhus Airport is located in Aarhus, Denmark. Aarhus Airport is situated 25.0 meters above sea level. Aarhus Airport has a runway called 10R/28L.
Aarhus_Airport is located in Aarhus, Denmark. Aarhus_Airport is situated 25.0 meters above sea level. Aarhus_Airport has a runway called 10R/28L.

\[ \downarrow \text{form} \]

NAME_{S1} is located in NAME_{O2}. PRONOUN_{S1} is situated NAME_{O3} meters above sea level. DESCRIPTION_{S1} has a runway called NAME_{O4}. 
FERREIRA

Surface Realization

$NAME_{S1}$ is located in $NAME_{o2}$. $PRONOUN_{S1}$ is situated $NAME_{o3}$ meters above sea level.
$DESCRIPTION_{S1}$ has a runway called $NAME_{o5}$.

$\downarrow realize$

Pick the most frequent referring expression, given entity, form, syntactic position and referential status.
Features extracted from the dependency tree of the wikified text.
AUTOMATIC EVALUATION

REG metrics
Accuracy, string edit distance and pronoun accuracy

Text metrics
Text accuracy and BLEU
## REG METRICS

<table>
<thead>
<tr>
<th></th>
<th>Acc</th>
<th>String</th>
<th>Pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only Names</td>
<td>53%$^D$</td>
<td>4.05$^D$</td>
<td>-</td>
</tr>
<tr>
<td>Ferreira</td>
<td>61%$^C$</td>
<td>3.18$^C$</td>
<td>43%$^B$</td>
</tr>
<tr>
<td>NeuralREG+Seq2Seq</td>
<td>74%$^{A,B}$</td>
<td>2.32$^{A,B}$</td>
<td>75%$^A$</td>
</tr>
<tr>
<td>NeuralREG+CAtt</td>
<td>74%$^A$</td>
<td>2.25$^A$</td>
<td>75%$^A$</td>
</tr>
<tr>
<td>NeuralREG+HierAtt</td>
<td>73%$^B$</td>
<td>2.36$^B$</td>
<td>73%$^A$</td>
</tr>
<tr>
<td>Model</td>
<td>Acc</td>
<td>BLEU</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>-----</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Only Names</td>
<td>15%</td>
<td>69.03</td>
<td></td>
</tr>
<tr>
<td>Ferreira</td>
<td>19%</td>
<td>72.78</td>
<td></td>
</tr>
<tr>
<td>NeuralREG+Seq2Seq</td>
<td>28%</td>
<td>79.27</td>
<td></td>
</tr>
<tr>
<td>NeuralREG+CAtt</td>
<td>30%</td>
<td>79.39</td>
<td></td>
</tr>
<tr>
<td>NeuralREG+HierAtt</td>
<td>28%</td>
<td>79.01</td>
<td></td>
</tr>
</tbody>
</table>
HUMAN EVALUATION

Material
144 trials (6 triple set sizes × 4 instances × 6 text versions)

Method
Latin square design
24 trials/list (144 trials ÷ 6 lists)
60 participants (10 participants/list)

Metrics
Fluency, Grammaticality and Clarity
7-Likert scale
## HUMAN EVALUATION

<table>
<thead>
<tr>
<th>Method</th>
<th>Fluency</th>
<th>Grammar</th>
<th>Clarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only Names</td>
<td>4.74(^C)</td>
<td>4.68(^B)</td>
<td>4.90(^B)</td>
</tr>
<tr>
<td>Ferreira</td>
<td>4.74(^C)</td>
<td>4.58(^B)</td>
<td>4.93(^B)</td>
</tr>
<tr>
<td>NeuralREG+Seq2Seq</td>
<td>4.95(^B,C)</td>
<td>4.82(^A,B)</td>
<td>4.97(^B)</td>
</tr>
<tr>
<td>NeuralREG+CAtt</td>
<td>5.23(^A,B)</td>
<td>4.95(^A,B)</td>
<td>5.26(^A,B)</td>
</tr>
<tr>
<td>NeuralREG+HierAtt</td>
<td>5.07(^B,C)</td>
<td>4.90(^A,B)</td>
<td>5.13(^A,B)</td>
</tr>
<tr>
<td>Original</td>
<td>5.41(^A)</td>
<td>5.17(^A)</td>
<td>5.42(^A)</td>
</tr>
</tbody>
</table>
CONCLUSION

First end-to-end approach for REG in text discourse
Improvements over reference accuracy and text fluency
Concatenative attention (CAtt) best decoding method

Delexicalized version of WebNLG corpus
Useful resource for NLG in general

Data and code available
https://github.com/ThiagoCF05/NeuralREG
<table>
<thead>
<tr>
<th>Model</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>OnlyNames</td>
<td>Alan Shepard was born in New Hampshire on 1923-11-18. Before Alan Shepard death in California Alan Shepard had been awarded Distinguished Service Medal (United States Navy) an award higher than Department of Commerce Gold Medal.</td>
</tr>
<tr>
<td>Ferreira</td>
<td>Alan Shepard was born in New Hampshire on 1923-11-18. Before his death in California him had been awarded Distinguished Service Medal an award higher than Department of Commerce Gold Medal.</td>
</tr>
<tr>
<td>Seq2Seq</td>
<td>Alan Shepard was born in New Hampshire on 1923-11-18. Before his death in California him had been awarded the Distinguished Service Medal by the United States Navy an award higher than the Department of Commerce Gold Medal.</td>
</tr>
<tr>
<td>CAtt</td>
<td>Alan Shepard was born in New Hampshire on 1923-11-18. Before his death in California he had been awarded the Distinguished Service Medal by the US Navy an award higher than the Department of Commerce Gold Medal.</td>
</tr>
<tr>
<td>HierAtt</td>
<td>Alan Shephard was born in New Hampshire on 18 November 1923. Before his death in California he had been awarded the Distinguished Service Medal an award higher than the Department of Commerce Gold Medal.</td>
</tr>
<tr>
<td>Original</td>
<td>Alan Shepard was born in New Hampshire on 18 November 1923. Before his death in California he had been awarded the Distinguished Service Medal by the US Navy an award higher than the Department of Commerce Gold Medal.</td>
</tr>
</tbody>
</table>

Thank you! :-)
https://github.com/ThiagoCF05/NeuralREG
## SETTINGS

<table>
<thead>
<tr>
<th>Setting</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Method</td>
<td>Adam</td>
</tr>
<tr>
<td>Matrices init</td>
<td>Xavier</td>
</tr>
<tr>
<td>Batch Size</td>
<td>40</td>
</tr>
<tr>
<td>Epochs</td>
<td>60</td>
</tr>
<tr>
<td>Embedding Size</td>
<td>300</td>
</tr>
<tr>
<td>Hidden Layer Size</td>
<td>512</td>
</tr>
<tr>
<td>Dropout</td>
<td>0.2/0.3</td>
</tr>
</tbody>
</table>