Building Adaptable and Scalable Natural Language Generation Systems

Yannis Konstas
The president of the United States Donald Trump announced that he would not go to the annual dinner of the White House Correspondents' Association (WHCA) in late April.
The president of the United States Donald Trump announced that he would not go to the annual dinner of the White House Correspondents’ Association (WHCA) in late April.
Natural Language Generation is everywhere (Dialogue Systems)

Here's the forecast for Sitka, AK for today:

Sitka, AK

PM T-Storms
Chance of Rain: 30%
High: 4° Low: -1°

1°
Natural Language Generation is everywhere
(Dialogue Systems)
Natural Language Generation is everywhere (Dialogue Systems)

Here's the forecast for Sitka, AK for today:

Sitka, AK
PM T-Storms
Chance of Rain: 30%
High: 4° Low: -1°

Here's the forecast for Chicago, IL for today:

Chicago, IL
Partly Cloudy
Chance of Rain: 10%
High: 19° Low: 8°

Don't forget your raincoat if you're going to Chicago, IL, Yannis...

How about on Tuesday

Chicago, IL
Tuesday, 28 Feb
Scattered Showers
Chance of Rain: 60%
14°
Low: 6°
Natural Language Generation is everywhere (Conversational Agents)

...or when things get too emotional
Natural Language Generation is everywhere (Educational Technology)

(Harsley et al., CSCW 2016)
Natural Language Generation is everywhere
(Caption Generation)

A man swinging a bat.

(Krause et al., CVPR 2017)
Natural Language Generation is everywhere
(Caption Generation)

A baseball player is swinging a bat.
He is wearing a red helmet and a white shirt.
The catcher’s mitt is behind the batter.

(Krause et al., CVPR 2017)
Machine Translation
Text Summarization
Code to Language
Dialogue Systems
Instructional Text
Meaning Representations
Educational Technology
Concept-to-Text
Human-Robot Interaction
Conversational Agents
Storytelling
Captions
Natural Language Generation

Input: Computer-interpretable representation of the world

- **Select** content
- **Organize** content in particular order
- Decide how to **verbalise** content

Output: Text
High quality source code is often paired with high level summaries of the computation it performs, for example in code documentation or in descriptions posted in online forums.

```csharp
public int TextWidth(string text) {
    TextBlock t = new TextBlock();
    t.Text = text;
    return (int)Math.Ceiling(t.ActualWidth);
}
```

<table>
<thead>
<tr>
<th></th>
<th>min</th>
<th>mean</th>
<th>max</th>
<th>mod</th>
</tr>
</thead>
<tbody>
<tr>
<td>wind</td>
<td>10</td>
<td>15</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>dir</td>
<td></td>
<td></td>
<td>W</td>
<td></td>
</tr>
<tr>
<td>temp</td>
<td>50</td>
<td>60</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td>gust</td>
<td>5</td>
<td>10</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

Meaning Representations

- know
- inhabit
- I
- planet
- man
- lazy

Educational Technology

20x + 5y = γ

Human-Robot Interaction

Machine Translation

Concept-to-Text

Code to Language

Human-Robot Interaction
I know the planet is inhabited by a lazy man.

Tammy bought 20 apples and 5 oranges. How many fruits does she have now?

Place the Heineken block west of the Mercedes block.

Overcast, with a high of 70. Moderate westerly winds, with gusts as high as 13 mph.

High quality source code is often paired with high level summaries of the computation it performs, for example in code documentation or in descriptions posted in online forums.

Get rendered width of string rounded up to the nearest integer.

```csharp
public int TextWidth (string text) {
    TextBlock t = new TextBlock();
    t.Text = text;
    return (int)Math.Ceiling(t.ActualWidth);
}
```

Machine Translation

Meaning Representations

Concept-to-Text

Educational Technology
Existing Approaches

Successes

- Rule-based frameworks
- Modular architecture
Existing Approaches

Successes

‣ Rule-based frameworks
‣ Modular architecture

Challenges

‣ **Expensive** to build
‣ **Hard to deploy** to new applications
Data-driven NLG

- **Learn** generation process *directly* from data
- **Easier** to build and maintain
- **Adapt** to multiple domains
Data-driven NLG

- **Learn** generation process *directly* from data
- **Easier** to build and maintain
- **Adapt** to multiple domains

Challenges

- **Require** large corpora - NLG is **low-resourced**
- **New** machine learning model for **every** application
Outline

- **Neural Network** architecture for NLG
  - **Learn** from different inputs
Outline

- **Neural Network** architecture for NLG
  - Learn from different inputs

- **Address** low-resource problem
  - Generic framework for **scaling** to large corpora without extra annotation
  - **Collect** large datasets from community-based platform
Outline

- **Neural Network** architecture for NLG
  - Learn from different inputs

- **Address** low-resource problem
  - Generic framework for *scaling* to large corpora without extra annotation
  - **Collect** large datasets from community-based platform

- **Adapt** to two applications
  - Meaning Representations
  - Code to Language
Neural NLG

Joint work with
Srinivasan Iyer, Mark Yatskar
Luke Zettlemoyer, Yejin Choi
Overview

- Sequence to sequence architecture
  - End-to-end model w/o intermediate representations
  - Linearisation of input to string
  - Pre-process

- Paired Training
  - Scalable data augmentation
Meaning Representations

Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

I knew a planet that was inhabited by a lazy man.
I have known a planet that was inhabited by a lazy man.
I know a planet. It is inhabited by a lazy man.

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)
Meaning Representations

Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

I knew a planet that was inhabited by a lazy man.
I have known a planet that was inhabited by a lazy man.
I know a planet. It is inhabited by a lazy man.

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)
Meaning Representations

Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

I knew a planet that was inhabited by a lazy man.

I have known a planet that was inhabited by a lazy man.

I know a planet. It is inhabited by a lazy man.

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)
Meaning Representations

Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

I knew a planet that was inhabited by a lazy man.

I have known a planet that was inhabited by a lazy man.

I know a planet. It is inhabited by a lazy man.

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)
Sequence to sequence model
Sequence to sequence model

\[ \hat{w} = \arg\max_w \prod_i p(w_i|w_{<i}, h^{(s)}) \]
Sequence to sequence model

\[ \hat{w} = \arg\max_w \prod_i p(w_i | w_{<i}, h^{(s)}) \]
US officials held an expert group meeting in January 2002 in New York.
US officials held an expert group meeting in January 2002 in New York.
Encoding

Linearize —> RNN encoding

hold

:ARG0 (person
  :ARG0-of (have-role
    :ARG1 United_States
    :ARG2 official)
  )

:ARG1 (meet
  :ARG0 (person
    :ARG1-of expert
    :ARG2-of group)
  )

:time (date-entity 2002 1)
:location New_York
Encoding

Linearize —> RNN encoding

```
hold
:ARG0 (person
  :ARG0-of (have-role
    :ARG1 United_States
    :ARG2 official)
  )
:ARG1 (meet
  :ARG0 (person
    :ARG1-of expert
    :ARG2-of group)
  )
:time (date-entity 2002 1)
:location New_York
```
Encoding

Linearize —> RNN encoding
- Token embeddings

hold
  :ARG0 (person
       :ARG0-of (have-role
                 :ARG1 United_States
                 :ARG2 official)
  )
  :ARG1 (meet
       :ARG0 (person
             :ARG1-of expert
             :ARG2-of group)
  )
  :time (date-entity 2002 1)
  :location New_York
Encoding

Linearize —> RNN encoding
- Token embeddings
- Recurrent Neural Network (RNN)

```
hold
  :ARG0 (person
    :ARG0-of (have-role
      :ARG1 United_States
      :ARG2 official)
  )
  :ARG1 (meet
    :ARG0 (person
      :ARG1-of expert
      :ARG2-of group)
  )
  :time (date-entity 2002 1)
  :location New_York
```
Encoding

Linearize —> RNN encoding
- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold
    :ARG0 (person
        :ARG0-of (have-role
            :ARG1 United_States
            :ARG2 official)
    )
    :ARG1 (meet
        :ARG0 (person
            :ARG1-of expert
            :ARG2-of group)
    )
    :time (date-entity 2002 1)
    :location New_York
```
Encoding

Linearize —> RNN encoding
- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold :ARG0 (person :ARG0-of (have-role :ARG1 United_States :ARG2 official))
   :ARG1 (meet :ARG0 (person :ARG1-of expert :ARG2-of group))
   :time (date-entity 2002 1)
   :location New_York
```
Encoding

Linearize —> RNN encoding
- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold
  :ARG0 (person
    :ARG0-of (have-role
      :ARG1 United_States
      :ARG2 official)
    )
  :ARG1 (meet
    :ARG0 (person
      :ARG0-of expert
      :ARG2-of group)
    )
  :time (date-entity 2002 1)
  :location New_York
```
Decoding

RNN Encoding $\rightarrow$ RNN Decoding (Beam search)
Decoding

RNN Encoding $\rightarrow$ RNN Decoding (Beam search)

- init $h^{(s)}$
Decoding

RNN Encoding —> RNN Decoding (Beam search)

- init $h^{(s)}$
- softmax

Holding
Held
US
...

$h_{N(s)}$ —> $h_1$
Decoding

RNN Encoding —> RNN Decoding (Beam search)

- init $h^{(s)}$
- softmax
- $p(w_i|w_{<i}, h^{(s)})$

Holding
Held
US
...

$w_{11}$: Holding
$w_{12}$: Helds
$w_{13}$: Hold
$w_{14}$: US
...

RNN Encoding —> RNN Decoding (Beam search)
Decoding

RNN Encoding —> RNN Decoding (Beam search)

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

RNN Encoding —> RNN Decoding (Beam search)

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$

- $\text{init } h^{(s)}$
- $\text{softmax}$
- $p(w_i|w_{<i}, h^{(s)})$
Decoding

RNN Encoding → RNN Decoding (Beam search)

- init $h^{(s)}$
- softmax
- $p(w_i|w_{<i}, h^{(s)})$

Holding
Held
US
the
meeting
US
person
expert
meeting
meetings
meet

$w_{11}$: Holding
$w_{12}$: Helds
$w_{13}$: Hold
$w_{14}$: US
... 

$w_{21}$: Hold a
$w_{22}$: Hold the
$w_{23}$: Held a
$w_{24}$: Held the
... 

$w_{k1}$: The US officials held
$w_{k2}$: US officials held a
$w_{k3}$: US officials hold the
$w_{k4}$: US officials will hold a
...
Attention

\( h_2 \) \( \rightarrow \) \( h_3 \)

\( \text{a the meeting} \)

\( w_2: \text{held} \)
Attention

hold → [h1(s)]
ARG0 → [h2(s)]
(kwargs) → [h2(s)]
person → [h3(s)]
ARG0-of → [h4(s)]

w2: held

h3 → c3

w2: held

[a the meeting …]
Attention

\[ w_2: \text{held} \]

\[ a \text{ the meeting} \ldots \]

\[ c_3 \]

\[ \begin{bmatrix} \end{bmatrix} \]

\[ \mathbf{a}_i = \text{soft} \max \left( \mathbf{f}_i \left( \mathbf{h}^{(s)}, \mathbf{h}_i \right) \right) \]

\[ \mathbf{c}_i = \sum_i \mathbf{a}_{ij} \mathbf{h}_j^{(s)} \]
Attentive language model

\[ w_2: \text{hold} \]

\[ \text{the meeting} \]

\[ a \]

\[ \text{hold} \]

\[ \text{ARG0} \]

\[ ( \]

\[ \text{person} \]

\[ \text{ARG0-of} \]

\[ a_i = \text{softmax}(f_i(h^{(s)}, h_i)) \]

\[ c_i = \sum a_{ij} h_j^{(s)} \]

US officials held an expert group meeting in January 2002
US officials held an expert group meeting in January 2002.
Attention

US officials held an expert group meeting in January 2002.

\[ a_i = \text{softmax} \left( f_i \left( h^{(s)}, h_i \right) \right) \]

\[ c_i = \sum_i a_{ij} h_j^{(s)} \]
US officials held an expert group meeting in January 2002.
US officials held an expert group meeting in January 2002 in New York.
US officials held an expert group meeting in January 2002 in New York.
US officials held an expert group meeting in January 2002 in New York.
US officials held an expert group meeting in January 2002 in New York.
US officials held an expert group meeting in January 2002 in New York.

loc_0 officials held an expert group meeting in month_0 year_0 in loc_1.
Experimental Setup

AMR LDC2015E86 (SemEval-2016 Task 8)

- Hand annotated MR graphs: newswire, forums
- \(\sim 16k\) training / 1k development / 1k test pairs

Train

- Optimize cross-entropy loss

Evaluation

- BLEU n-gram precision
  (Papineni et al., ACL 2002)
First Attempt

**TreeToStr**: Flanigan et al, NAACL 2016

**TSP**: Song et al, EMNLP 2016

**PBMT**: Pourdamaghani and Knight, INLG 2016
First Attempt

<table>
<thead>
<tr>
<th>Method</th>
<th>BLEU</th>
</tr>
</thead>
<tbody>
<tr>
<td>TreeToStr</td>
<td>23</td>
</tr>
<tr>
<td>TSP</td>
<td>22.4</td>
</tr>
<tr>
<td>PBMT</td>
<td>26.9</td>
</tr>
</tbody>
</table>

**TreeToStr**: Flanigan et al, NAACL 2016
**TSP**: Song et al, EMNLP 2016
**PBMT**: Pourdamaghani and Knight, INLG 2016
First Attempt

TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016
First Attempt

All systems use a Language Model trained on a very large corpus. We will emulate via data augmentation.

TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

(Sennrich et al., ACL 2016)
What went wrong?

Reference

US officials held an expert group meeting in January 2002 in New York.

Prediction

United States officials held held a meeting in January 2002.
What went wrong?

Reference
US officials held an expert group meeting in January 2002 in New York.

Prediction
United States officials held a meeting in January 2002.

- Repetition

hold
:ARG0 (person
  :ARG0-of (have-role
    :ARG1 loc_0
    :ARG2 official)
  )
:ARG1 (meet
  :ARG0 (person
    :ARG1-of expert
    :ARG2-of group)
  )
:time (date-entity year_0 month_0)
:location loc_1
US officials held an expert group meeting in January 2002 in New York.

United States officials held a meeting in January 2002.

- Repetition
- Coverage
What went wrong?

Reference
US officials held an expert group meeting in January 2002 in New York.

Prediction
United States officials held a meeting in January 2002.

- Repetition
- Coverage
  a) Sparsity

Tokens

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>OOV@1</th>
<th>OOV@5</th>
</tr>
</thead>
<tbody>
<tr>
<td>OOV@1</td>
<td>44.26%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OOV@5</td>
<td></td>
<td></td>
<td>74.85%</td>
</tr>
</tbody>
</table>
What went wrong?

US officials held an expert group meeting in January 2002 in New York.

United States officials held a meeting in January 2002.

- Repetition
- Coverage
  a) Sparsity
  b) Avg sent length: 20 words
  c) Limited Language Modeling capacity
Data Augmentation

Original Dataset: ~16k graph-sentence pairs
Data Augmentation

Original Dataset: ~16k graph-sentence pairs

Gigaword: ~183M sentences *only*
Data Augmentation

Original Dataset: ~16k graph-sentence pairs

Gigaword: ~183M sentences *only*

Sample sentences with vocabulary overlap
Data Augmentation

Generate from MR

Encoder → Decoder

Attention
Data Augmentation

Parse to MR

Generate from MR

text → Encoder → Decoder → graph

graph → Encoder → Decoder → text

Attention

graph
Data Augmentation

Parse to MR

Generate from MR
Data Augmentation

- Parse to MR
- Generate from MR
- Re-train
Data Augmentation

Parse to Input

input

text

Generate from Input
Paired Training
Paired Training

Train MR Parser \( P \) on Original Dataset
Paired Training

Train MR Parser $P$ on Original Dataset

for $i = 0 \ldots N$
Paired Training

Train MR Parser $\mathbf{P}$ on Original Dataset

For $i = 0 \ldots N$

$S_i =$ Sample $k 10^i$ sentences from Gigaword
Paired Training

Train MR Parser $\mathbf{P}$ on Original Dataset

\[
\text{for } i = 0 \ldots N
\]

$S_i =$ Sample $k \times 10^i$ sentences from Gigaword

Parse $S_i$ sentences with $\mathbf{P}$
Paired Training

Train MR Parser $P$ on Original Dataset

```
for i = 0 … N

$S_i$ = Sample $k \cdot 10^i$ sentences from Gigaword
Parse $S_i$ sentences with $P$
Re-train MR Parser $P$ on $S_i$
```
**Paired Training**

- Train MR Parser $P$ on Original Dataset

```plaintext
for i = 0 … N

$S_i =$Sample $k \times 10^i$ sentences from Gigaword
Parse $S_i$ sentences with $P$
Re-train MR Parser $P$ on $S_i$
```
Paired Training

Train MR Parser $P$ on Original Dataset

for $i = 0 \ldots N$

$S_i =$ Sample $k \times 10^i$ sentences from Gigaword

Parse $S_i$ sentences with $P$

Re-train MR Parser $P$ on $S_i$

Train Generator $G$ on $S_N$
Training MR Parser

Train $P$ on Original Dataset
Training MR Parser

Train P on Original Dataset
Training MR Parser

Train $P$ on Original Dataset

Sample $S_1=200k$ sentences from Gigaword

[Diagram showing a document being parsed into an AMR representation]

200k
Training MR Parser

Train P on Original Dataset

Sample $S_1=200k$ sentences from Gigaword

Parse $S_1$ with P
Training MR Parser

Train P on Original Dataset

Sample $S_1=200k$ sentences from Gigaword

Parse $S_1$ with P

Train P on $S_1=200k$
Training MR Parser

- Train $P$ on Original Dataset
- Sample $S_1 = 200k$ sentences from Gigaword
- Parse $S_1$ with $P$
- Fine-tune $P$ on Original Dataset

200k
Training MR Parser

Sample $S_2=2M$ sentences from Gigaword

Fine-tune $P$ on Original Dataset

Parse $S_2$ with $P$ ( )

Train $P$ on $S_2=2M$

Fine-tune: init parameters from previous step and train on Original Dataset
Training MR Parser

- Sample $S_2 = 2M$ sentences from Gigaword
- Fine-tune $P$ on Original Dataset
- Parse $S_2$ with $P$
- Train $P$ on $S_2 = 2M$

*Fine-tune:* init parameters from previous step and train on Original Dataset.
Training MR Generator

Sample $S_3=2M$ sentences from Gigaword

Parse $S_3$ with $P$ ( )

Fine-tune $G$ on Original Dataset

Train $G$ on $S_3=2M$

Fine-tune: init parameters from previous step and train on Original Dataset
Training MR Generator

Sample $S_3=2M$ sentences from Gigaword

Parse $S_3$ with $P$

Fine-tune $G$ on Original Dataset

Train $G$ on $S_3=2M$
Training MR Generator

Sample $S_3 = 2M$ sentences from Gigaword

Train $G$ on $S_3 = 2M$

Fine-tune $G$ on Original Dataset

Parse $S_3$ with $P$

Fine-tune: init parameters from previous step and train on Original Dataset
Final Results

**TreeToStr**: Flanigan et al, NAACL 2016
**TSP**: Song et al, EMNLP 2016
**PBMT**: Pourdamaghani and Knight, INLG 2016
Final Results

TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016
Final Results

TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016
Final Results

TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016
Final Results

**TreeToStr**: Flanigan et al, NAACL 2016

**TSP**: Song et al, EMNLP 2016

**PBMT**: Pourdamaghani and Knight, INLG 2016
How did we do?

Reference

US officials held an expert group meeting in January 2002 in New York.

Prediction

In January 2002 United States officials held a meeting of the group experts in New York.

Errors: Disfluency Coverage
How did we do?

Reference
US officials held an expert group meeting **in January 2002** in New York.

Prediction
**In January 2002** United States officials held a meeting of the group experts in New York.

Reference
The report stated **British government** must help to stabilize **weak states** and push for international regulations that would stop **terrorists** using freely available information to create and unleash new forms of biological warfare such as **a modified** version of the influenza **virus**.

Prediction
The report stated that the **Britain government** must help stabilize **the weak states** and push international regulations to stop the use of freely available information to create a form of new biological warfare such as **the modified** version of the influenza **virus**.

Errors: **Disfluency Coverage**
Adapt to other applications?

- **Structured** input representation
  - ✔️ Meaning Representation of Natural Language
  - ☐ Programming Language
Code to Language

Joint work with

Srinivasan Iyer
Luke Zettlemoyer, Alvin Cheung
Code to Language

Input: Source Code
(SQL - C#)

```
public int TextWidth (string text) {
    TextBlock t = new TextBlock();
    t.Text = text;
    return (int) Math.Ceiling(t.ActualWidth);
}
```

Output: Summary
Get rendered width of string rounded up to the nearest integer.

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
Input: Source Code
(SQL - C#)

```csharp
public int TextWidth (string text) {
    TextBlock t = new TextBlock();
    t.Text = text;
    return (int) Math.Ceiling(t.ActualWidth);
}
```

Output: Summary
Get rendered width of string rounded up to the nearest integer.

```
SELECT max(marks) FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

How to find the second largest value from a table?

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
Input Representation

1) Code snippet —> Linearize (left-to-right)

```sql
SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

How to find the second largest value from a table?

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
1) Code snippet —> Linearize (left-to-right)

2) Anonymize

SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);

SELECT max(col0)
FROM tab0
WHERE col0 < (SELECT max(col1) FROM tab1);

How to find the second largest value from a table?
Input Representation

1) Code snippet —> Linearize (left-to-right)
2) Anonymize
3) Bag of Words Encoding

How to find the second largest value from a table?

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
Decoding with Attention

4) Bag of Words Encoding —> RNN Decoding
5) Attention directly on input embeddings

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
Decoding with Attention

4) Bag of Words Encoding —> RNN Decoding
5) Attention directly on input embeddings

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)
Community-based Datasets
Community-based Datasets

How to find the Second largest value from a table?

One table with EmpSalary in Employee Table. I need to find the second largest Salary what is paid by the company.

3

How to find the Second largest value(Salary) from a table.

sql-server-2005 tsql

share improve this question edited Oct 18 '10 at 10:56

5 What value do you want returned if there are two records with the equal top value? – Oct 18 '10 at 10:48

add a comment

3 Answers

Try this: this should give the second largest salary:

3

SELECT MAX(EmpSalary) FROM employee WHERE EmpSalary < (SELECT MAX(EmpSalary) FROM employee);

share improve this answer answered Oct 18 '10 at 10:59
Community-based Datasets

- (Accepted Answer, Post title) pairs
- ~33K SQL / 66k C# examples
Results

**PBMT**: MOSES Phrase-based MT system

**SUM-NN**: Rush et al, EMNLP 2015
Results

PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015
Results

PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015
Human Evaluation Results

**Naturalness**

- **IR**
- **PBMT**
- **SUM-NN**
- **CODE-NN**

<table>
<thead>
<tr>
<th>System</th>
<th>SQL</th>
<th>C#</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBMT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM-NN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE-NN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Informativeness**

<table>
<thead>
<tr>
<th>System</th>
<th>SQL</th>
<th>C#</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBMT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUM-NN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CODE-NN</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**PBMT**: MOSES Phrase-based MT system

**SUM-NN**: Rush et al, EMNLP 2015
How did we do?

SELECT * FROM table
ORDER BY Rand() LIMIT 10

Reference
Select random rows from mysql table

CODE-NN
How to get random rows from a mysql database?
How did we do?

SELECT * FROM table
ORDER BY Rand() LIMIT 10

How to get random rows from a mysql database?

Adding childs to a treenode dynamically in C#?

How to get all child nodes in TreeView?
Neural NLG Contributions
Neural NLG Contributions

- **Adapt** to multiple applications
- **Scale** to very large corpora

- Address low-resource problem
  - **Paired training** general technique
  - Train on **noisy** community-based datasets
Future Work
Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

Joint work with

Rik Koncel-Kedziorski
Luke Zettlemoyer, Hannaneh Hajishirzi

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

Syntactic, Semantic, Thematic rewriter

Joint work with

Rik Koncel-Kedzierski
Luke Zettlemoyer, Hannaneh Hajishirzi

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

= Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

\[ 504 + x = 639 \]
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

\[504 + x = 639\]
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

\[ 504 + x = 639 \]
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

\[504 + x = 639\]
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?
Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.
Educational Technology

Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

504 + x = 639

504 + x = 639

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?
Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?
Educational Technology

Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

Defense lawyer Thomas Olsson stated it was very tragic and a failure for Swedish law and order that the client Thomas Olsson was representing had been kept in detention. The official alleged Karzai was reluctant to move against big drug lords in Karzai’s political power base.

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

504 + x = 639
Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

Bob has 639 sheep. Alice has 504 sheep. How many more sheep does Bob have than Alice?

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?
Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)
Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Mei et al., NAACL 2016)
### Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Mei et al., NAACL 2016)
Caption Generation

(Krause et al., CVPR 2017, Yatskar et al., CVPR 2016, Krishna et al., 2016)
A baseball player is **swinging** a bat. He is **wearing** a red helmet and a white shirt. The catcher’s mitt is **behind** the batter.

(Krause et al., CVPR 2017, Yatskar et al., CVPR 2016, Krishna et al., 2016)
Semantic-based Machine Translation

Source

The children told that lie

Target

そのうそは子供たちがついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

(Jones et al., COLING 2012)
Semantic-based Machine Translation

Source

The children told that lie

Target

そのうそは子供たちがついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

(Jones et al., COLING 2012)
Semantic-based Machine Translation

Source: The children told that lie

Target: そのうそは子供たちがついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST
Semantic-based Machine Translation

The children told that lie

Source

Graph-to-graph transformation

Target

そのうそは子供たちがついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

(Jones et al., COLING 2012)
Semantic-based Machine Translation

The children told that lie.

そのうそは子供たちがついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

(Jones et al., COLING 2012)
Semantic-based Machine Translation

Graph-to-graph transformation

Source

The children told that lie

Target

そのうそは子供たちがついた
sono use-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

› No Japanese AMR corpus

(Jones et al., COLING 2012)
Semantic-based Machine Translation

Graph-to-graph transformation

Source: The children told that lie

Target: そのうそは 子供たちが ついた
sono uso-wa kodomo-tachi-ga tsui-ta
that lie-TOP child-and others-NOM breathe out-PAST

- No Japanese AMR corpus
- MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)

Joint work with Michael Wayne Goodman

(Jones et al., COLING 2012)
Semantic-based Machine Translation

- No Japanese AMR corpus
- MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)

1) **Parse to MRS** from English

*(Jones et al., COLING 2012)*
Semantic-based Machine Translation

The children told that lie

Source

No Japanese AMR corpus
MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)
1) Parse to MRS from English
2) Generate Japanese from MRS

Target

Joint work with Michael Wayne Goodman

(Jones et al., COLING 2012)
Dialogue Systems

(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)
> I would like to follow up on my speech therapy treatment.
Dialogue Systems

> I would like to follow up on my speech therapy treatment.

(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)
Dialogue Systems

> I would like to follow up on my speech therapy treatment.

(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)
I would like to follow up on my speech therapy treatment.

Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. [...] Verbal impairment related to communication impairment was treated with speech therapy 3 months ago. [...]
I would like to follow up on my speech therapy treatment.

Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. [...] Verbal impairment related to communication impairment was treated with speech therapy 3 months ago. [...]
I would like to follow up on my speech therapy treatment.

Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. [...] Verbal impairment related to communication impairment was treated with speech therapy 3 months ago. [...]
I would like to follow up on my speech therapy treatment.  

I can see in my logs, that we started improving verbal impairment due to the accident, with speech therapy 3 months ago. When would you like to book the next appointment?
Summary

- General data-driven approach for NLG
- Facilitates the deployment to new domains
- Integrates to existing systems and applications
Summary

- General data-driven approach for NLG
- Facilitates the deployment to new domains
- Integrates to existing systems and applications

Thank You