It’s not just how you say it, it’s what you say: Improving response generation in conversational agents

Marie Meteer
Brandeis University
Computational Linguistics
Dialog Response Generation

Dialog Act Selection
(usually in the "dialog manager")

Predefined strings

Templates

[Diagram showing the flow from Dialog Act Selection to Templates to Predefined strings]
How can we make progress?
Early Hybrid approaches: Apply a statistical language model

Statistical Ngram Language Model

Corpus

Score

Template based text generation

Version 1
Version 2
Version 3
Version 4

Best one
Early Hybrid approaches: Addressing fluency with a statistical language model

- **Phrasal pattern:**
  - flying FROMCITY FROMTIME TOCITY TOTIME

- **No n-gram weights**
  - flying from Newark at 7:20am to Boston arriving in Boston at 8:00am
  - flying Newark at 7:20am to Boston that arrive at 8:00am

- **With n-gram weights**
  - flying from Newark at 7:20am arriving in Boston at 8:00am
  - flying from Newark leaving at 7:20am and arriving in Boston at 8:00am

Response generation as Information Retrieval

Find the closest statement

Find the response that best matches the statement

Return that response

Response generation as Summarization

Use the question as a query

Retrieve a matching document

Select “summary” sentence(s)

Question

Query

Document

Response?
Response generation as Translation

- **English sentence**
- **French sentence**
- **Parallel Eng/Fr corpus**
- **Q & A corpus**
- **Twitter comment**
- **Twitter response**

**English Question**

**English Answer**

**Your favorite learning algorithm**
What did the AlexPrize bots do?

Waterfall:
- Rules/Templates
  - Retrieval (IR or Summary)
  - Generative (with selection)

Hybrid:
- Generative
  - Rules/Templates
  - Retrieval

Where does it fail?

The most likely response is not always the best one.

A: Where are you going?
B: I’m going to the restroom.
A: See you later.
B: See you later.
A: See you later.
B: See you later.
A: …

Dialog context
Didn’t you just say that?

A: What’s your name?
B: Daniel.
A: How old are you?
B: Twelve. What’s your name?
A: Peter.
B: How old are you?
A: Eleven. What’s your name?
B: Daniel.
A: How old are you?
B: …
What do we need for fluent NLG?

How to use the “little words”

<table>
<thead>
<tr>
<th>Discourse context</th>
<th>Application context</th>
<th>The world</th>
</tr>
</thead>
<tbody>
<tr>
<td>He/him</td>
<td>Both</td>
<td>Lunch</td>
</tr>
<tr>
<td>Still</td>
<td>The other one</td>
<td>?</td>
</tr>
<tr>
<td>Again</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which ones …</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It’s not just applying machine learning, it’s figuring out what the data and features should be.

(as Jay said this morning)
What are we missing?
Back to the Future with CwC

- DARPA’s *Communicating with Computers* goes back to SHRDLU’s Blocks World

- Complex ideas are composed from elementary ones;
- Most elementary ideas are about the physical world;
- Language specifies how to compose complex ideas; but,
- Context is often needed to boost the specificity of complex ideas that can be composed given language


This work was supported by DARPA Contract W911NF-15-C-0238.
The challenge for you all
(and especially the machine learning experts)

Make what’s in your world a part of the learning

Keep an eye on the robots
Need help?

• Brandeis University
  – Computational Linguistics MS
  – Summer “JBS” on Building Mobile Voice Applications
  – Contact Marie Meteer
    • mmeteer@brandeis.edu