Data-Driven Response Generation

Alan Ritter
Ohio State University
1950s ~ 2010
Dialog systems mostly rule-based

Chatbots:
- Rule-Based: Eliza (Weizenbaum 1966)
- Information Retrieval (Isbell et. al. 2000)

Goal-Directed Dialogue Systems:
- ATIS Dataset (Hemphill, 1990)
  - 774 flight reservation conversations
  - Manually annotated

On my way to the airport
Have a safe flight!
Thanks!
1990s ~ 2010s
Data-Driven Machine Translation
millions of bilingual documents on the web

Findings of WMT 2010 (Callison-Burch et. al.)
The Mathematics of Statistical Machine Translation: Parameter Estimation (Brown et. al.)
July 2011
Data-Driven Dialogue
500 million conversations per month on Twitter alone
(vs. 30m for French-English translation)

On my way to the airport
Have a safe flight!
Thanks!

Alan Ritter, Colin Cherry, Bill Dolan (EMNLP 2011) “Data-Driven Response Generation in Social Media”
July 2011

Data-Driven Dialogue

500 million conversations **per month** on Twitter alone

(vs. 30m for French-English translation)

**NLP on Noisy User-Generated Text:**

- Unsupervised Dialogue Acts
  (Ritter, Cherry, Dolan, NAACL 2010)

- Named Entity Recognition
  (Ritter et. al. EMNLP 2011)

- Open-Domain Event Extraction
  (Ritter et. al. KDD 2012)

- Minimally-Supervised Event Extraction
  (Ritter et. al. WWW 2015)

Alan Ritter, Colin Cherry, Bill Dolan (EMNLP 2011) “Data-Driven Response Generation in Social Media”
July 2011
Data-Driven Dialogue
500 million conversations per month on Twitter alone

(vs. 30m for French-English translation)

... and they lived happily ever after.

Alan Ritter, Colin Cherry, Bill Dolan (EMNLP 2011) "Data-Driven Response Generation in Social Media"
But, unlike MT, conversations are not semantically equivalent.
Who wants to come over for dinner tomorrow?
Input:
Who wants to come over for dinner tomorrow?

Output:
Yum! I
Who wants to come over for dinner tomorrow?

Yum! I want to
Input:

Who wants to come over for dinner tomorrow?

Output:

Yum! I want to be there
Who wants to come over for dinner tomorrow?

Input:

Output:

Yum! I want to be there tomorrow!
2015 ~ present
Neural MT-based Conversation Models

- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao and Bill Dolan. A Diversity-Promoting Objective Function for Neural Conversation Models. NAACL 2016

Alan Ritter, Colin Cherry, Bill Dolan (EMNLP 2011) “Data-Driven Response Generation in Social Media”
But, maximum likelihood estimate responses can be safe and boring

$$\arg\max_{r_1, \ldots, r_l} P(r_1, \ldots, r_l|m_1, \ldots, m_k)$$

Response

Input Message
But, maximum likelihood estimate responses can be safe and boring

$$\arg \max_{r_1, \ldots, r_l} P(r_1, \ldots, r_l | m_1, \ldots, m_k)$$

Some replies work for almost any input:
But, maximum likelihood estimate responses can be safe and boring

$$\arg \max_{r_1, \ldots, r_l} P(r_1, \ldots, r_l|m_1, \ldots, m_k)$$

Some replies work for almost any input:

“I don’t know”
2016
Neural Dialogue with Deep Reinforcement Learning

Neural Dialogue with Deep Reinforcement Learning

Jiwei Li (PhD Stanford 2017)
Problem: Short-sighted conversation decisions.

Problem: Short-sighted conversation decisions.

Problem: Short-sighted conversation decisions.

How old are you?

16?

i'm 16.
Problem: Short-sighted conversation decisions.

How old are you?

16?

i'm 16.

i don't know what you're talking about

Problem: Short-sighted conversation decisions.

How old are you?

i'm 16.

16?

i don't know what you're talking about.

Problem: Short-sighted conversation decisions.

How old are you?

i'm 16.

16?

i don't know what you're talking about.

you don't know what you're saying.
Problem: Short-sighted conversation decisions.

How old are you?

16?

you don't know what you're saying

i'm 16.

i don't know what you're talking about

i don't know what you're talking about
Problem: Short-sighted conversation decisions.

How old are you?

16?

you don’t know what you’re saying

you don’t know what you’re saying

i’m 16.

i don’t know what you’re talking about

i don’t know what you’re talking about
How old are you?

16?

you don’t know what you’re saying

you don’t know what you’re saying

Bad Action

i’m 16.

i don’t know what you’re talking about

i don’t know what you’re talking about

Problem: Short-sighted conversation decisions.

Problem: Short-sighted conversation decisions.

- How old are you?
  - I'm 16.
- 16?
  - I don't know what you're talking about
  - You don't know what you're saying
  - I don't know what you're talking about
- Outcome
Can Reinforcement Learning Handle This?

How old are you?

i'm 16.

16?

i don't know what you're talking about

you don't know what you're saying

i don't know what you're talking about

you don't know what you're saying

Outcome does not emerge until a few turns later

Notations: State

$\mathbf{r}_{i-1}$

How old are you?

Encoding

- how
- old
- are
- you

How old are you?

I'm 16.

Encoding:
- how
- old
- are
- you

Decoding:
- I'm
- 16
- fine
- EOS

Notations: Action
A message from training set

Encode

r1

Decode

r2

...
Compute Accumulated Reward $R(S_1, S_2, \ldots, S_n)$
Policy Gradient Methods:
REINFORCE Algorithm (William, 1992)

\[ J(\theta) = \mathbb{E}[R(s_1, s_2, \ldots, s_N)] \]
Policy Gradient Methods:
REINFORCE Algorithm (William, 1992)

\[ J(\theta) = \mathbb{E}[R(s_1, s_2, \ldots, s_N)] \]

\[ \nabla J(\theta) = \nabla \log p(s_1, s_2, \ldots, s_N)R(s_1, s_2, \ldots, s_N) \]

What we want to learn

Q: How to specify a reward signal?

A: Turing Test
Q: How to specify a reward signal?

A: Turing Test

Adversarial Learning
(Goodfellow et al., 2014)
Adversarial Learning for Neural Dialogue

Real-world conversations

Response Generator

sample human response

Discriminator

Real or Fake?

generate response

Adversarial Learning for Neural Dialogue

Adversarial Learning for Neural Dialogue

Real-world conversations

(Alternate Between Training Generator and Discriminator)

sample human response

Discriminator

Real or Fake?

generate response

Response Generator

REINFORCE Algorithm (Williams, 1992)

Adversarial Learning Improves Response Generation

Human Evaluator:

Machine Evaluator:

<table>
<thead>
<tr>
<th></th>
<th>Adversarial Win</th>
<th>Adversarial Lose</th>
<th>Tie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adversarial</td>
<td>62%</td>
<td>18%</td>
<td>20%</td>
</tr>
<tr>
<td>Adversarial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(How often can</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>you fool a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>machine)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adversarial Learning Improves Response Generation

vs a vanilla generation model

Adversarial Learning

- **8.0%**

Standard Seq2Seq model

- **4.9%**

Future: Integrating dynamic knowledge graphs

Extract Entities, Relations and Events

Takeaways
Takeaways

Open-Domain Dialogue

Dialogue  MT
Takeaways

Open-Domain Dialogue

Learning from Delayed-Reward
Takeaways

Open-Domain Dialogue

Dialogue

MT

Learning from Delayed-Reward

Adversarial Learning for Dialogue

Alan Ritter (Ohio State University)
Takeaways

Open-Domain Dialogue

Dialogue

MT

Learning from Delayed-Reward

Adversarial Learning for Dialogue

Thank You!