Deep Learning in Open Domain Dialogue Generation

Jiwei Li
Computer Science Department
Stanford University
May 10, 2017
• The voice-driven personal assistant on your iPhone
• Perhaps the most visible & exciting application of NLU today
• A major breakthrough in artificial intelligence (AI)??
• The next generation of interaction design??

Borrowed From Bill MacCartney’s slides
Does Siri really understand language?
Does Siri really understand language?
Does Siri really understand language?

*Colbert: Write the show.*
Does Siri really understand language?

Colbert: Write the show.

Siri: What would you like to search for?
Does Siri really understand language?

**Colbert:** Write the show.

**Siri:** What would you like to search for?

**Colbert:** I don’t want to search for anything! I want to write the show!
Does Siri really understand language?

Colbert: Write the show.

Siri: What would you like to search for?

Colbert: I don’t want to search for anything! I want to write the show!

Siri: Searching the Web for “search for anything. I want to write the shuffle.”
Does Siri really understand language?

Colbert: Write the show.

Siri: What would you like to search for?

Colbert: I don’t want to search for anything! I want to write the show!

Siri: Searching the Web for “search for anything. I want to write the shuffle.”

[...]

Borrowed From Bill MaĐCaƌtŶey's slides
Does Siri really understand language?

Colbert: Write the show.
Siri: What would you like to search for?
Colbert: I don’t want to search for anything! I want to write the show!
Siri: Searching the Web for “search for anything. I want to write the shuffle.”

[...]
Colbert: For the love of God, the cameras are on, give me something?
Does Siri really understand language?

Colbert: Write the show.
Siri: What would you like to search for?
Colbert: I don’t want to search for anything! I want to write the show!
Siri: Searching the Web for “search for anything. I want to write the shuffle.”
[...]
Colbert: For the love of God, the cameras are on, give me something?
Siri: What kind of place are you looking for? Camera stores or churches?
Reminiscent of Eliza (1966)

**User:** You are like my father in some ways.

**ELIZA:** What resemblance do you see

**User:** You are not very aggressive but I think you don’t want me to notice that.

**ELIZA:** What makes you think I am not aggressive?

**User:** You don’t argue with me.

**ELIZA:** Why do you think I don’t argue with you?

**User:** You are afraid of me.

**ELIZA:** Does it please you to believe I am afraid of you?
How Eliza works

Key word list

If a keyword is identified

manipulate the input

Else:

give a generic response, or copy some utterance from dialogue history
Background

(Ritter et al., 2010,
Sordoni, et al., 2015,
Vinyals and Le, 2015)
Background

(Ritter et al., 2010, Sordoni, et al., 2015, Vinyals and Le, 2015)

Goal Oriented Tasks
Why is building a chatbot hard?

1. Computers need to understand what you ask.
Why is building a chatbot hard?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask,
Why is building a chatbot hard?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask, that require domain knowledge, discourse knowledge, world knowledge.
Why is building a chatbot hard?

1. Computers need to understand what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask, that require domain knowledge, discourse knowledge, world knowledge.
3. Responses from the same bot need to be consistent.
Why is building a chatbot hard?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask, that require **domain knowledge**, **discourse knowledge**, **world knowledge**
3. Responses from the same bot need to be consistent.
4. A bot should be interactive.
Background

1. IR based model
2. Generation models
A big conversation corpus

A: How old are you
B: I am eight

A: What’s your name?
B: I am John

A: How do you like CS224n?
B: I cannot hate it more.

A: How do you like Jiwei?
B: He’s such a Jerk !!!!!
IR based model

A big conversation corpus

A: How old are you
B: I am eight

A: What’s your name?
B: I am john

A: How do you like CS224n?
B: I cannot hate it more.

A: How do you like Jiwei?
B: He’s such a Jerk !!!!!

An new input:

What’s your age?
IR based model

A big conversation corpus

A: How old are you
B: I am eight

A: What’s your name?
B: I am John

A: How do you like CS224n?
B: I cannot hate it more.

A: How do you like Jiwei?
B: He’s such a Jerk !!!!!
IR based model

A big conversation corpus

A: How old are you
B: I am eight

A: What's your name?
B: I am john

A: How do you like CS224n?
B: I cannot hate it more.

A: How do you like Jiwei?
B: He's such a Jerk !!!!!

An new input:

What’s your age?
IR based model

A big conversation corpus

A: How old are you
B: I am eight

A: What’s your name?
B: I am john

A: How do you like CS224n?
B: I cannot hate it more.

A: How do you like Jiwei?
B: He’s such a Jerk !!!!!

An new input:

What’s your age?
I am eight.
Generation Models
Generation Models

Response Generation as Statistical Machine Translation (Ritter et al., 2010)

Slide borrowed from Michel Galley
Response Generation as SMT

Exploit high-frequency patterns with phrase-based MT

“I am” → “you are”  “sick” → “get better”  “lovely!” → “thanks!”

(Ritter et al., 2010)

Slide borrowed from Michel Galley
Seq2Seq Models for Response Generation
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

\[ \text{Loss} = - \log p(\text{target}|\text{source}) \]

Source: Input Messages
Target: Responses

how are you? I'm fine. EOS

Encoding Decoding

Source: Input Messages
Target: Responses
Seq2Seq Models for Response Generation

how are you?
Seq2Seq Models for Response Generation

Encoding

how are you ?
Seq2Seq Models for Response Generation

Encoding

how are you?
Seq2Seq Models for Response Generation

how are you?
Seq2Seq Models for Response Generation

Encoding

how are you ?
Seq2Seq Models for Response Generation

Encoding

how

are

you

?

Decoding
Seq2Seq Models for Response Generation

how are you?
I’m
Encoding Decoding
eos
Seq2Seq Models for Response Generation

how are you? I'm fine

Encoding Decoding

eos I'm
Seq2Seq Models for Response Generation

Encoding

Decoding

how are you ?

I’m fine.
Seq2Seq Models for Response Generation

how are you?

I'm fine.

Encoding

Decoding
<table>
<thead>
<tr>
<th>Input</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>what is moral ?</td>
<td>what empowered humanity</td>
</tr>
<tr>
<td>what is immoral ?</td>
<td>the fact that you have a child .</td>
</tr>
<tr>
<td>what is the purpose of existence ?</td>
<td>to find out what happens when we get to the planet earth .</td>
</tr>
<tr>
<td>what do you think about bill gates ?</td>
<td>He’s a good man</td>
</tr>
</tbody>
</table>
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Generic Responses
Generic Responses

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015)
Generic Responses

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015; )
Generic Responses

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015;)

How is life?

I don’t know what you are talking about.
How to deal with Generic Responses

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015;)

30% percent of all generated responses
How to deal with Generic Responses

```python
Def ChatBot(string):
    if string[len(string)-1] == "?":
        return "I don’t know"
    else:
        return "I don’t know what you are talking about"
```
How to deal with Generic Responses

```python
Def ChatBot(string):
    if string[len(string)-1] == "?":
        return "I don’t know"
    else:
        return "I don’t know what you are talking about"

What Eliza is doing !!
How to deal with Generic Responses

Solution #1: Adding Rules
How to deal with Generic Responses

Solution #1: Adding Rules

I don’t know.
I don’t know..
I don’t know ...
...
I don’t know !
I don’t know !!
I don’t know !!!
How to deal with Generic Responses

Solution #1: Adding Rules

I don’t know.
I don’t know..
I don’t know ...
...
I don’t know !
I don’t know !!
I don’t know !!!

I have no idea.
I don’t have a clue.
I don’t have the foggiest idea what you are talking about.
I don’t have the lightest idea what you are talking about.
I haven’t the faintest idea

How should I know?
How to deal with Generic Responses

Solution #1: Adding Rules

I don’t know.
I don’t know..
I don’t know ...
... 
I don’t know!
I don’t know!!
I don’t know!!!

I have no idea.
I don’t have a clue.
I don’t have the lightest idea what you are talking about.
I don’t have the foggiest idea what you are talking about.
I haven’t the faintest idea

How should I know?

Rules don’t work!!
Mutual Information for Response Generation.

\[
    \text{Loss} = - \log p(\text{target}|\text{source})
\]
Mutual Information for Response Generation.

$$\text{Loss} = - \log p(\text{target}|\text{source})$$

$p(\text{target}|\text{source})$

source $\rightarrow$ target
Mutual Information for Response Generation.

\[ \text{Loss} = - \log p(\text{target}|\text{source}) \]

“I don’t know”

Whatever one asks
Mutual Information for Response Generation.

$p(target|source)$

source $\rightarrow$ target

What one asks $\rightarrow$ “I don’t know”?
Mutual Information for Response Generation.

What one asks → "I don’t know"

"I don’t know" → What one asks
Mutual Information for Response Generation.
Mutual Information for Response Generation.

\[
\text{Mutual Information} = \log \frac{p(S, T)}{p(S)p(T)}
\]
Mutual Information for Response Generation.

\[ \hat{T} = \arg \max_T \left\{ \log \frac{p(S, T)}{p(S)p(T)} \right\} \]
Mutual Information for Response Generation.

\[
\hat{T} = \arg \max_T \left\{ \log p(T|S) - \lambda \log p(T) \right\}
\]

Bayesian Rule

\[
T = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}
\]
Mutual Information for Response Generation.

\[
\hat{T} = \arg \max_T \left\{ \log p(T|S) - \lambda \log p(T) \right\}
\]

\[
T = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}
\]

Bayesian Rule

\[ p(\text{target}|\text{source}) \]
Mutual Information for Response Generation.

\[
\hat{T} = \arg \max_T \left\{ \log p(T|S) - \lambda \log p(T) \right\}
\]

Bayesian Rule

\[
T = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}
\]

\[p(\text{source}|\text{target})\]
Sampled Results

| Input: What are you doing? | Standard Seq2Seq p(t|s) | Mutual Information |
|----------------------------|-------------------------|--------------------|
| -0.86 I don’t know.        | -1.09 Get out of here.  |                    |
| -1.03 I don’t know!        | -1.09 I’m going home.   |                    |
| -1.06 Nothing.             | -1.09 Oh my god!        |                    |
| -1.09 Get out of the way.  | -1.10 I’m talking to you.|                    |

<table>
<thead>
<tr>
<th>Input: what is your name?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.91 I don’t know.</td>
</tr>
<tr>
<td>-0.92 I don’t know!</td>
</tr>
<tr>
<td>-0.92 I don’t know, sir.</td>
</tr>
<tr>
<td>-0.97 Oh, my god!</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: How old are you?</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.79 I don’t know.</td>
</tr>
<tr>
<td>-1.06 I’m fine.</td>
</tr>
<tr>
<td>-1.17 I’m all right.</td>
</tr>
<tr>
<td>-1.17 I’m not sure.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input: What are you doing?</th>
<th>1. I’ve been looking for you.</th>
<th>2. I want to talk to you.</th>
<th>3. Just making sure you’re OK.</th>
<th>4. I told you to shut up.</th>
<th>5. Get out of here.</th>
<th>6. I’m looking for a doctor.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Twenty-eight.</td>
<td>3. Twenty-four.</td>
<td>4. Five.</td>
<td>5. 15.</td>
<td></td>
<td>6. Eight.</td>
<td></td>
</tr>
<tr>
<td>3. Long.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Multi-context Response Generation

Single Context:

Any particular plan?

????
Multi-context Response Generation

What's your plan for the upcoming summer?

I am going to Hawaii for vacation.

Any particular plan?

????
What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

Any particular plan?

????
Multi-context Response Generation

Notations

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

Any particular plan?

Response $r$

????
Multi-context Response Generation

Notations

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

Any particular plan?

Response r

Message: m

????
What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

Any particular plan?

Response r

????

Message: m
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

...

Any particular plan?
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vacation.

... ...

Any particular plan?

LSTM
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

... ...

Any particular plan?

$c_i = \frac{h_i^T \cdot h_m}{\sum_k \exp(h_i^T h_m)}$

$H$
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

\[ C = \sum_i O_i c_i \]
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.

Weighted Sum

Memory Network (Weston et al., 2014)

\[
C = \sum_i O_i c_i
\]
Multi-context Response Generation

where do you live

Encoding

Decoding

\[ C = \sum_{i} O_{ic_i} \]
Multi-context Response Generation

where do you live

Encoding

Decoding

Attention Models (Bahdanau et al., 2014; Luong et al., 2015)

\[ O_i = \frac{\exp(c_i^T \cdot h_{t-1})}{\sum_i \exp(c_i^T \cdot h_{t-1})} \]

\[ C = \sum_i O_i c_i \]
Multi-context Response Generation

\[ O_i = \frac{\exp(c_i^T \cdot h_{t-1})}{\sum_i \exp(c_i^T \cdot h_{t-1})} \]

\[ C = \sum_i O_i c_i \]
Multi-context Response Generation

Encoding

where
do
you
live

Decoding

in

EOS

C

in

C

\[ O_i = \frac{\exp(c_i^T \cdot h_{t-1})}{\sum_i \exp(c_i^T \cdot h_{t-1})} \]

\[ C = \sum_i O_i c_i \]
Multi-context Response Generation

where do you live

Encoding

Decoding

\[ O_i = \frac{\exp (c_i^T \cdot h_{t-1})}{\sum_i \exp (c_i^T \cdot h_{t-1})} \]

\[ C = \sum_i O_i c_i \]
Multi-context Response Generation

where do you live

Encoding

$m$

Decoding

in

uk

. 

EOS

\[
O_i = \frac{\exp(c_i^T \cdot h_{t-1})}{\sum_i \exp(c_i^T \cdot h_{t-1})}
\]

\[
C = \sum_i O_i c_i
\]
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Speaker Consistency
Speaker Consistency

I’m 8.

How old are you?
Speaker Consistency

How old are you?

I’m 8.

What’s your age?

18
Speaker Consistency

I live in Los Angeles.

Where do you live now?

I live in Los Angeles.
Speaker Consistency

Where do you live now?

I live in Los Angeles.

In which city do you live now?

I live in Paris.
I live in Los Angeles.

I live in Paris.

England, you?
Speaker Consistency

How old are you?

I’m 8.
Speaker Consistency

How old are you?

I’m 8.

How many kids do you have?

4, you?
When were you born?

In 1942.
Speaker Consistency

When were you born?
In 1942.

When was your mother born?
In 1966.
How to represent users

Persona embeddings (70k)
How to represent users

Persona embeddings (70k)
Bob

Word embeddings (50k)
uk london
great sydney
good
stay

monday
tuesday

live okay

stay
Persona seq2seq model

Encoding

where

do

you

live

Decoding

EOS
Persona seq2seq model

Encoding

Decoding

where do you live

Encoding

Decoding

 EOS

Persona embeddings (70k)

Bob
Persona seq2seq model

Encoding

where -> do -> you -> live

Decoding

EOS -> Bob

Persona embeddings (70k)
Persona seq2seq model

where do you live

Encoding

Decoding

EOS

Persona embeddings (70k)
Persona seq2seq model

where do you live
Encoding
Decoding
EOS
Persona embeddings (70k)
Bob
in
uk
Bob
Persona seq2seq model

Encoding

Decoding

where do you live

Bob in uk

EO

Word embeddings (50k)

Bob
great
good
london
sydney

stay

monday
tuesday

Bob

Word embeddings (50k)
Interaction Seq2Seq model

Encoding

where do you live

where  do  you  live
Interaction Seq2Seq model

where do you live

Encoding

\[ \text{tanh}(W* \ \cdots) \]
Interaction Seq2Seq model

where do you live

Encoding
\[ \text{tanh}(W^* \text{input}) \]

Decoding
\[ \text{EOS} \]

in

\[ \text{tanh}(W^* \text{output}) \]
Interaction Seq2Seq model

where do you live

Encoding

Decoding

tanh(W* )
in

in

uk

in

tanh(W* )

in

uk
Interaction Seq2Seq model

where do you live

Encoding

Decoding

in

uk

in

uk

EOS

where

do

you

live

in

uk

.
<table>
<thead>
<tr>
<th>User1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>message</td>
<td>Where is your hometown?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I’m from England.</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>Where are you from?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I’m from England.</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>In which city do you live now?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I live in London.</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>In which country do you live now?</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>What is your major?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>Business. You?</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>What did you study in college?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I did business studies.</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>How old are you?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I’m 18.</td>
<td></td>
</tr>
<tr>
<td>message</td>
<td>What is your age?</td>
<td></td>
</tr>
<tr>
<td>response</td>
<td>I’m 18.</td>
<td></td>
</tr>
<tr>
<td>User1</td>
<td>message</td>
<td>response</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>Where is your hometown?</td>
<td>I'm from <strong>England</strong>.</td>
</tr>
<tr>
<td></td>
<td>Where are you from?</td>
<td>I'm from <strong>England</strong>.</td>
</tr>
<tr>
<td></td>
<td>In which city do you live now?</td>
<td>I live in <strong>London</strong>.</td>
</tr>
<tr>
<td></td>
<td>In which country do you live now?</td>
<td>I live in <strong>England</strong>.</td>
</tr>
<tr>
<td></td>
<td>What is your major?</td>
<td>Business. You?</td>
</tr>
<tr>
<td></td>
<td>What did you study in college?</td>
<td>I did business studies.</td>
</tr>
<tr>
<td></td>
<td>How old are you?</td>
<td>I'm 18.</td>
</tr>
<tr>
<td></td>
<td>What is your age?</td>
<td>I'm 18.</td>
</tr>
<tr>
<td>User1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>Where is your hometown?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I’m from England.</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>Where are you from?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I’m from England.</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>In which city do you live now?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I live in London.</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>In which country do you live now?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I live in England.</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What is your major?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td><strong>Business</strong></td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>You?</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What did you study in college?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I did <strong>business studies</strong></td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>How old are you?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I’m 18.</td>
<td></td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What is your age?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I’m 18.</td>
<td></td>
</tr>
</tbody>
</table>
Results (No cherry-picking)

User1

message Where is your hometown?
response I’m from England.
message Where are you from?
response I’m from England.
message In which city do you live now?
response I live in London.
message In which country do you live now?
message What is your major?
response Business. You?
message What did you study in college?
response I did business studies.
message How old are you?
response I’m 18.
message What is your age?
response I’m 18.

User2

message Where is your hometown?
response I live in Indonesia.
message Where are you from?
response England, you?
message In which city do you live now?
response I live in Jakarta.
message In which country do you live now?
response I live in Indonesia.
message What is your major?
response Business, you?
message What did you study in college?
response Psychology, you?
message How old are you?
response I’m 18.
message What is your age?
response I’m 16.
Issues

How do we handle long-term dialogue success?
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Issues

How do we handle long-term dialogue success?

• Problem 1: Dull and generic responses.
Issues

• Problem 1: Dull and generic responses.

“I don’t know” problem (Sordoni et al., 2015; Serban et al., 2015;)

Do you love me?

I don’t know what you are talking about.
Issues

How do we handle long-term dialogue success?

• Problem 1: Dull and generic responses.
• Problem 2: Repetitive responses.
Problem 2: Repetitive responses.

Shut up!
Problem 2: Repetitive responses.

Shut up!

No, you shut up!
Problem 2: Repetitive responses.

Shut up!

No, you shut up!

No, you shut up!
Problem 2: Repetitive responses.
Problem 2: Repetitive responses.

Shut up!
No, you shut up!
No, you shut up!
No, you shut up!
Problem 2: Repetitive responses.

See you later!

See you later!
Issues

How do we handle long-term dialogue success?

• Problem 1: Dull and generic responses.
• Problem 2: Repetitive responses.
• Problem 3: Short-sighted conversation decisions.
Problem 3: Short-sighted conversation decisions.

How old are you?
Problem 3: Short-sighted conversation decisions.

How old are you?

I'm 16.
Problem 3: Short-sighted conversation decisions.

How old are you?

16?

i'm 16.
Problem 3: Short-sighted conversation decisions.

How old are you?

I'm 16.

16?

I don't know what you're talking about.
Problem 3: Short-sighted conversation decisions.

How old are you?

i'm 16.

16?

i don't know what you're talking about
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 3: Short-sighted conversation decisions.
Problem 3: 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
Problem 3: 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 don't know what you're talking about

I don't know what you're talking about

I'm 16.
Problem 3: 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.

Bad Action
Problem 3: Short-sighted conversation decisions.

How old are you?

16?

i don't know what you're talking about

i don't know what you're saying

Outcome

Problem 3: Short-sighted conversation decisions.
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

you don't know what you're saying

i don't know what you're talking about

Outcome does not emerge until a few turns later
Can reinforcement learning handle this?
Notations for Reinforcement Learning
Notations: **State**

$r_{i-1}$

**Encoding**

How old are you?

how old are you

how    old    are    you
Notations: Action

$r_{i-1}$

How old are you?

$r_{i}$

i'm 16.
Notations: **Reward**
Notations: **Reward**

$r_{i-1}$

How old are you?

$r_i$

i'm 16.
Notations: **Reward**

1. Ease of answering
Notations: **Reward**

1. Ease of answering

\[ R_1 = - \sum_{s \in S} \log p_{\text{seq2seq}}(s | a) \]
Notations: **Reward**

1. Ease of answering

\[ R_1 = - \sum_{s \in S} \log p_{\text{seq2seq}}(s|a) \]

S: ”I don’t know what you are talking about”
Notations: Reward

2. Information Flow
Notations: **Reward**

2. Information Flow

See you later!
Notations: **Reward**

2. Information Flow

\[ r_2 = -\log \text{Sigmoid}(\cos(s_1, s_2)) \]
Notations: **Reward**

3. Meaningfulness

\[
\log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1)
\]

S1: How old are you?

S2: I'm 16.
Notations: **Reward**

1. Easy to answer        R1
2. Information Flow      R2
3. Meaningfulness        R3

\[ R = \lambda_1 R_1 + \lambda_2 R_2 + \lambda_3 R_3 \]
Simulation

A message from training set
Simulation

A message from training set

Encode
Simulation

A message from training set

Encode

Decode

r1

...
Simulation

A message from training set → Encode → Decode → r1 → Encode...

...
Simulation

A message from training set

Encode

r1

Decode

r2

…
Input Message

Encode Decode

Turn 1
Encode

Turn 2
Decode Encode

…
Decode

Turn N
S1 S2 Sn

Compute Accumulated Reward R(S1,S2,…,Sn)
1. Easy to answer

\[ Pr(\text{Dull Response}|S_1) \quad Pr(\text{Dull Response}|S_2) \quad Pr(\text{Dull Response}|S_n) \]
1. Easy to answer  
2. Information Flow  

Input 
Message 
Encode Decode 

Turn 1 
Encode 

Turn 2 
Decode Encode 

… 
Decode 

Turn N 
S1 S2 S3 

$-\log\text{Sigmoid} \cos(s_1, s_2)$
1. Easy to answer R1
2. Information Flow R2
3. Meaningfulness R3

Input Message
Encode
Turn 1
Encode
Turn 2
Decode Encode
…
Decode
Turn N

\[ \log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1) \]
1. Easy to answer  R1
2. Information Flow    R2
3. Meaningfulness      R3

Input 
Message 
Encode Decode 

Turn 1 
Encode 

Turn 2 
Decode Encode 

… 
Decode 

Turn N 

S1 S2 S3

\[
\log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1)
\]

\[
\log p_{\text{seq2seq}}(s_2|s_3) + \log p_{\text{seq2seq}}(s_3|s_2)
\]
1. Easy to answer  
2. Information Flow  
3. Meaningfulness  

Input Message

Encode

Decode

Turn 1

Encode

Turn 2

Decode Encode

…

Decode

Turn N

S1 S2 S3

\[
\log p_{\text{seq2seq}}(s_1 | s_2) + \log p_{\text{seq2seq}}(s_2 | s_1)
\]

\[
\log p_{\text{seq2seq}}(s_2 | s_3) + \log p_{\text{seq2seq}}(s_3 | s_2)
\]
1. Easy to answer \( R_1 \)
2. Information Flow \( R_2 \)
3. Meaningfulness \( R_3 \)

Input
Message
Encode
Decode

Turn 1
Encode

Turn 2
Decode
Encode

…
Decode
Encode

Turn \( N \)

S_1 \ S_2 \ S_3

Compute Accumulated Reward \( R(S_1, S_2, \ldots, S_n) \)

\[ \log p_{\text{seq2seq}}(s_1 | s_2) + \log p_{\text{seq2seq}}(s_2 | s_1) \]

\[ \log p_{\text{seq2seq}}(s_2 | s_3) + \log p_{\text{seq2seq}}(s_3 | s_2) \]
Input Message

Encode Decode

Turn 1
Encode
Turn 2
Decode Encode

... Decode
Turn N
S1 S2 Sn

REINFORCE Algorithm (William, 1992)

\[ J(\theta) = \mathbb{E}[R(s_1, s_2, ..., s_N)] \]
REINFORCE Algorithm (William, 1992)

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

\[ \nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) R(s_1, s_2, ..., s_N) \]
REINFORCE Algorithm (William, 1992)

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

\[ \nabla J(\theta) = \nabla \log p(s_1, s_2, \ldots, s_N) R(s_1, s_2, \ldots, s_N) \]
REINFORCE Algorithm (William, 1992)

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

\[ \nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) R(s_1, s_2, ..., s_N) \]
REINFORCE Algorithm (William, 1992)

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

\[
\nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) R(s_1, s_2, ..., s_N)
\]
REINFORCE Algorithm (William, 1992)

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

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

\[ \nabla J(\theta) = \nabla \log \prod_i p(s_i|s_{i-1}) R(s_1, s_2, \ldots, s_N) \]
What we want to learn

REINFORCE Algorithm (William, 1992)

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

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

\[ \nabla J(\theta) = \nabla \log \prod_i p(s_i|s_{i-1}) R(s_1, s_2, ..., s_N) \]
## Results

<table>
<thead>
<tr>
<th>Input</th>
<th>Mutual Information</th>
<th>The proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>How old are you ?</td>
<td>I’m 16.</td>
<td>I’m 16. why are you asking</td>
</tr>
</tbody>
</table>

**Input Mutual Information**

- **How old are you?**
  - I'm 16.

**The proposed model**

- I'm 16. why are you asking
## Results

<table>
<thead>
<tr>
<th>Input</th>
<th>Mutual Information</th>
<th>The proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>How old are you?</td>
<td>I’m 16.</td>
<td>I’m 16. why are you asking?</td>
</tr>
<tr>
<td>I’m 16.</td>
<td>I’m 16. why are you asking?</td>
<td></td>
</tr>
</tbody>
</table>
## Results

<table>
<thead>
<tr>
<th>Input</th>
<th>Mutual Information</th>
<th>The proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>How old are you?</td>
<td>I’m 16.</td>
<td>I’m 16. why are you asking?</td>
</tr>
<tr>
<td>what is your full name?</td>
<td>i have no idea</td>
<td>what 's yours?</td>
</tr>
<tr>
<td>I don’t want to go home</td>
<td>Really?</td>
<td>Why?</td>
</tr>
</tbody>
</table>
### Results

<table>
<thead>
<tr>
<th>Input</th>
<th>Mutual Information</th>
<th>The proposed model</th>
</tr>
</thead>
<tbody>
<tr>
<td>How old are you?</td>
<td>I’m 16.</td>
<td>I’m 16. why are you asking ?</td>
</tr>
<tr>
<td>what is your full name?</td>
<td>i have no idea</td>
<td>what 's yours ?</td>
</tr>
<tr>
<td>I don’t want to go home tonight.</td>
<td>Really ?</td>
<td>Why ?</td>
</tr>
<tr>
<td>Do you have any feelings for me ?</td>
<td>I don’t know what you are talking about.</td>
<td>Would I see you if I didn’t ?</td>
</tr>
</tbody>
</table>
Simulation

How old are you?
How old are you?

I'm 16, why are you asking?
How old are you?

I thought you were 12.

I'm 16, why are you asking?
Simulation

How old are you?

I thought you were 12.

I'm 16, why are you asking?

What made you think so?
How old are you?

I thought you were 12.

I don’t know what you are talking about.

i’m 16. why are you asking?

What made you think so?
Simulation

How old are you?

I thought you were 12.

I don’t know what you are talking about.

I’m 16, why are you asking?

What made you think so?

You don’t know what you are saying.
How old are you?
I thought you were 12.
I don’t know what you are talking about.
I don’t know what you are talking about.
i’m 16, why are you asking?
What made you think so?
You don’t know what you are saying.
How old are you?

I thought you were 12.

I don’t know what you are talking about.

I don’t know what you are saying.

What made you think so?

You don’t know what you are saying.

I’m 16, why are you asking?

Survive 4 turns!!
Outline
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Reward for Good Dialogue
Reward for Good Dialogue

How to evaluate open domain dialogue generation
Reward for Good Dialogue

How to evaluate open domain dialogue generation

Bleu ? Ppl ?
Reward for Good Dialogue

Turing Test
Reward for Good Dialogue

Turing Test

Input

How old are you?

I’m 25.
Reward for Good Dialogue

Turing Test

Input

How old are you?

I don’t know what you are talking about

I’m 25.

I don’t know what you are talking about
reward for good dialogue

Turing Test

How old are you?

I’m 25.

I don’t know what you are talking about

A human evaluator/judge
How old are you?

I don’t know what you are talking about.

I’m 25.

A human evaluator/judge
How old are you?
I don’t know what you are talking about.
I’m 25.

To expensive?
I don’t know what you are talking about.
How old are you?

I’m 25.

I don’t know what you are talking about.

Reward for Good Dialogue
Reward for Good Dialogue

How old are you?

I'm 25.

I don't know what you are talking about.

P= 90% human generated

P= 10% human generated
How old are you?

I do n't know what you are talking about

I ' m 25.

Reward for Good Dialogue

P = 90% human generated

I don ' t know what you are talking about

P = 10% human generated
Adversarial Learning in Image Generation (Goodfellow et al., 2014)

I do not know what you are talking about.

I’m 25.

P = 90% human generated

I don’t know what you are talking about.

P = 10% human generated
Adversarial Learning in Image Generation (Goodfellow et al., 2014)

P = 10% human generated

Discriminator

P = 90% human generated

P = 10% human generated
Adversarial Learning in Image Generation (Goodfellow et al., 2014)
Model Break Down

Generative Model (G)

Discriminative Model (D)
Model Break Down

Generative Model (G)

Discriminative Model (D)
Model Break Down

Generative Model (G)

how are you?

I’m fine. EOS

Encoding

Decoding

how are you? I’m fine.

eos I’m fine.
Model Break Down

Generative Model (G)

how → are → you → ? → I’m → fine → .

Encoding

Discriminative Model (D)

how → are → you → ? → eos → I’m → fine → .
Model Break Down

Generative Model (G)

Encoding

how → are → you → ?

Decoding

I’m → fine → .

Discriminative Model (D)

how → are → you → ?

eos → I’m → fine → .
Model Break Down

Generative Model (G)
how are you?
I'm fine.

Encoding

Decoding

Discriminative Model (D)
how are you?
I'm fine.
P = 90% human generated
Model Break Down

Generative Model (G)

I'm fine.

Discriminative Model (D)

how are you?

I'm fine.

P = 90% human generated Reward
Policy Gradient

Generative Model (G)

how are you?
I’m fine EOS

Encoding

Decoding

REINFORCE Algorithm (William, 1992)
Policy Gradient

Generative Model (G)

![Diagram of encoding and decoding process with tokens representing "how are you? I'm fine."

REINFORCE Algorithm (William, 1992)

\[ J = E[R(y)] \]
Policy Gradient

Generative Model (G)

REINFORCE Algorithm (William, 1992)

\[ J = E[R(y)] \]
\[ \nabla J = R(y) \nabla \log p(y) \]
Policy Gradient

Generative Model (G)

Encoding

Decoding

REINFORCE Algorithm (William, 1992)

\[ J = E[R(y)] \]
\[ \nabla J = R(y) \nabla \log p(y) \]
\[ = R(y) \sum_i \log p(y_i | x, y_1, \ldots, y_{i-1}) \]
Policy Gradient

Generative Model (G)

\[ \nabla J = R(y) \nabla \log p(y) \]

\[ = R(y) \sum_i \log p(y_i | x, y_1, \ldots, i-1) \]
Policy Gradient

Generative Model (G)

\[ \nabla J = R(y) \nabla \log p(y) \]

\[ = R(y) \sum_{i} \log p(y_i | x, y_1, ..., y_{i-1}) \]

Same reward to all tokens tokens
Policy Gradient

Input:  What’s your name
human:  I am John
machine: I don’t know
Policy Gradient

Input: What’s your name
human: I am John
machine: I don’t know

- REINFORCE assigns the same negative reward to all tokens [I, don’t, know]
Policy Gradient

Input: What’s your name
human: I am John
machine: I don’t know

- REINFORCE assigns the same negative reward to all tokens \([I, \ don’t, \ know]\)

Negative reward for “I”
Policy Gradient

Input: What’s your name
human: I am John
machine: I don’t know

- REINFORCE assigns the same negative reward to all tokens \([I, don’t, know]\)

Reward for Every Generation Step (REGS)
Policy Gradient

Monte Carlo Search

Figure from Yu et al (2016)
Policy Gradient

Generative Model (G)

\[ \sum_i R(y_i) \log p(y_i | x, y_1, \ldots, y_{i-1}) \]
Policy Gradient

Generative Model (G)

Difference

\[ \sum_i R(y_i) \log p(y_i | x, y_1, ..., i-1) \]

\[ R(y) \sum_i \log p(y_i | x, y_1, ..., i-1) \]

Encoding

Decoding

how are you ?

I'm fine

I'm fine.
Policy Gradient

Generative Model (G)

Encoding

Decoding

\[
\sum_i R(y_i) \log p(y_i|x, y_1, \ldots, y_{i-1})
\]

\[
\sum_i [R(y_i) - b(y_i)] \log p(y_i|x, y_1, \ldots, y_{i-1})
\]

Actor-critic RL
Adversarial Evaluation: "Adversarial Success"

How often a system can fool a computer into believing that its generated response was from a human.

How are you?
I think "P = 90% human generated"

The machine evaluator is fooled!!!

I don't know
Results: Adversarial Learning Improves Response Generation

Human Evaluator

Machine Evaluator

Adversarial Success
(How often can you fool a machine)

<table>
<thead>
<tr>
<th>Adversarial Learning</th>
<th>Adversarial Win</th>
<th>Adversarial Lose</th>
<th>Tie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adversarial Learning</td>
<td>8.0</td>
<td>18%</td>
<td>20%</td>
</tr>
<tr>
<td>Standard Seq2Seq model</td>
<td>4.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

vs a vanilla generation model
Adversarial Learning for Neural Dialogue Generation

For number of training iterations do
  For i=1,D-steps do
    Sample (X,Y) from real data
    Sample \( \hat{Y} \sim G(\cdot | X) \)
    Update \( D \) using \((X, Y)\) as positive examples and \((X, \hat{Y})\) as negative examples.
  End
  Update the Discriminator

For i=1,G-steps do
  Sample (X,Y) from real data
  Sample \( \hat{Y} \sim G(\cdot | X) \)
  Compute Reward \( r \) for \((X, \hat{Y})\) using \( D \).
  Update \( G \) on \((X, \hat{Y})\) using reward \( r \)
  Teacher-Forcing: Update \( G \) on \((X, Y)\)
End

Update the Generator

The discriminator forces the generator to produce correct responses
The previous RL model only perform better on multi-turn conversations

<table>
<thead>
<tr>
<th>Setting</th>
<th>adver-win</th>
<th>adver-lose</th>
<th>tie</th>
</tr>
</thead>
<tbody>
<tr>
<td>single-turn</td>
<td>0.62</td>
<td>0.18</td>
<td>0.20</td>
</tr>
<tr>
<td>multi-turn</td>
<td>0.72</td>
<td>0.10</td>
<td>0.18</td>
</tr>
</tbody>
</table>
Sample response

Tell me ... how long have you had this falling sickness?

<table>
<thead>
<tr>
<th>System</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla-MLE</td>
<td>I’m not a doctor.</td>
</tr>
<tr>
<td>Vanilla-Sample</td>
<td>Well everything you did was totally untrue.</td>
</tr>
<tr>
<td>REINFORCE</td>
<td>I don’t know how long it’s been.</td>
</tr>
<tr>
<td>REGS Monte Carlo (Reward for Every Generation Step)</td>
<td>A few months, I guess.</td>
</tr>
</tbody>
</table>
Outline

1. How to deal with generic responses
2. Consider more context
3. How to preserve Speaker Consistency
4. How to foster long-term success
5. What are the good rewards for dialogue generation
6. Building interactive bots
Introduction

Case 1

How do you like CS224s?
Introduction

Case 1

How do you like CS224s?

What is CS224n?
Introduction

Case 1

How do you like CS224s?

What is CS224n?

Is it a class about speech?
Introduction

What is cs224n?

Case 1

CS224n is a class about speech and dialogue

How do you like CS224s?
Introduction

Who is Hom Tanks?
Oh. Yeah. I like the class.

Case 1

How do you like CS224s?

CS224n is a class about speech and dialogue

Who is Hom Tanks?
Oh. Yeah. I like the class.
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?

How do you like CS224n?

UNK
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?

How do you like UNK?
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?

How do you like UNK?

Give an output anyway
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?

How do you like UNK?

softmax
Forward
Backward
Introduction

What will Current Chatbot Systems Do?

How do you like UNK?

softmax
Forward
Backward
output
I hate it. It’s really awful
Introduction

What will Current Chatbot Systems Do?

How do you like CS224s?

Searching the Web for “how do you like CS224n”
We need to the bot to ask questions !!
We Identify three classes of Questions:
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner

Which movie did Tom Hanks star in?

List, if possible, all movies that Tom Hanks made his appearance in.
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner
Text Clarification
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?
I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)
List, if possible, all movies that Tom Hanks made his appearance in.

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner
2. Knowledge Operation - query how to perform reasoning steps necessary to answer
Case 2: Knowledge Operation.

Knowledge Base:
- Tom Hanks directed Larry Crowne
- Tom Hanks starred Forrest Gump
- Robert Zemeckis directed Forrest Gump

which movie Did Tom Hanks appear in ?
Case 2: Knowledge Operation.

Knowledge Base:
- Tom Hanks directed Larry Crowne
- Tom Hanks starred Forrest Gump
- Robert Zemeckis directed Forrest Gump

Which movie did Tom Hanks appear in?
Case 2: Knowledge Operation.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
Case 2: Knowledge Operation.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?

S: Can you give me a hint?
Case 2: Knowledge Operation.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?

S: Can you give me a hint?

T: It is related to the kb fact that
Forrest Gump starred actors Tom Hanks
Robin Wright, Gary Sinise
Case 2: Knowledge Operation.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?

S: Can you give me a hint?

T: It is related to the kb fact that Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise

S: Does it have something to do with Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise
Case 2: Knowledge Operation.

T: Which movie did Tom Hanks star in?

S: Can you give me a hint?

T: It is related to the kb fact that
Forrest Gump starred actors Tom Hanks
Robin Wright, Gary Sinise

S: Does it have something to do with
Forrest Gump starred actors Tom Hanks,
Robin Wright, Gary Sinise

T: That's correct. (+)
Case 2: Knowledge Operation.

T: Which movie did Tom Hanks star in?
S: Can you give me a hint?
T: It is related to the kb fact that Forrest Gump starred actors Tom Hanks, Robin Wright, Gary Sinise.
S: Forrest Gump
T: That's correct. (+)
S: Forrest Gump
In what scenarios does a bot need to ask questions?  

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks  
kb: Forrest Gump starred actors Tom Hanks  
kb: Forrest Gump starred actors Sally Field  
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner
2. Knowledge Operation - query how to perform reasoning steps necessary to answer
3. Knowledge Acquisition: query to gain missing knowledge necessary to answer

How do you like Hom Tanks?
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner
2. Knowledge Operation - query how to perform reasoning steps necessary to answer
3. Knowledge Acquisition: query to gain missing knowledge necessary to answer
Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?

Not in the KB
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?

S: I don’t know. What’s the answer?
In what scenarios does a bot need to ask questions?

**Case 3: Knowledge Acquisition.**

```
kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
S: I don’t know. What’s the answer?
T: The answer is Forrest Gump.
```
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
S: I don’t know. What’s the answer?
T: The answer is Forrest Gump.
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
S: I don’t know. What’s the answer?
T: The answer is Forrest Gump.

... Other questions/ Other answers
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

In the context of knowledge acquisition, bots might ask questions to gather information that is not directly available from their knowledge base (KB).

- **Knowledge Base (KB):**
  - Larry Crowne directed by Tom Hanks
  - Forrest Gump starred actors Tom Hanks
  - Forrest Gump starred actors Sally Field
  - Forrest Gump directed by Robert Zemeckis

- **Conversation Example: (T = Teacher, S = Student)
  
  T: Which movie did Tom Hanks star in?
  S: I don’t know. What’s the answer?
  T: The answer is Forrest Gump.

... Other questions/ Other answers

T: Which movie did Tom Hanks star in?
In what scenarios does a bot need to ask questions?

Case 3: Knowledge Acquisition.

kb: Larry Crowne directed by Tom Hanks
kb: Forrest Gump starred actors Tom Hanks
kb: Forrest Gump starred actors Sally Field
kb: Forrest Gump directed by Robert Zemeckis

T: Which movie did Tom Hanks star in?
S: I don’t know. What’s the answer?
T: The answer is Forrest Gump.

... Other questions/ Other answers

T: Which movie did Tom Hanks star in?
S: Forrest Gump
three classes of Questions:

1. Text Clarification - query how to interpret text of dialog partner
2. Knowledge Operation - query how to perform reasoning steps necessary to answer
3. Knowledge Acquisition: query to gain missing knowledge necessary to answer
When shall a bot asks a question?

Shall I ask a question??
Which movie did Tom Hanks star in?
Which movie did Tom Hanks star in?

What do you mean?
Which movie did Tom Hanks star in?

What do you mean?

-CostAQ
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)  +1
Which *movie* did Tom Hanks *star* in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)   +1

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Larry Crowne
Which *movie* did Tom Hanks *star* in?

What do you mean?

I mean which film did Tom Hanks appear in.

Larry Crowne

That’s incorrect (+) -1
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Larry Crowne

That’s incorrect (+) -1

Reward: -1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That's correct (+) +1

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)  +1

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?
I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)

Larry Crowne

That’s incorrect (--) 

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?
I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)

Larry Crowne

That’s incorrect (--)
Which **movie** did Tom Hanks **star** in?

What do you mean?

I mean which **film** did Tom Hanks **appear** in.

*Forest Gump.*

That’s correct (+)

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?

I mean which film did Tom Hanks appear in.

Forest Gump.

That’s correct (+)

Reward: 1-CostAQ
Which movie did Tom Hanks star in?

What do you mean?
I mean which film did Tom Hanks appear in.
Forest Gump.

That’s correct (+)

Reward: +1

Reward: 1-CostAQ

<table>
<thead>
<tr>
<th>Final Answer Correct</th>
<th>Asking Question</th>
<th>Not asking Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - costₐq</td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>-1 - costₐq</td>
<td></td>
<td>-1</td>
</tr>
</tbody>
</table>
Policy Gradient

\[ \nabla J(\theta) = \nabla \log p_{RL}(a)[r - b] \]
Setting 2: Reinforcement Learning

T: Which **movie** did Tom Hanks **star** in?

Memory Network

Ask a question or not ..... 

<table>
<thead>
<tr>
<th>Final Answer Correct</th>
<th>Asking Question</th>
<th>Not asking Question</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$1 - \text{cost}_{AQ}$</td>
<td>1</td>
</tr>
<tr>
<td>Final Answer Incorrect</td>
<td>$-1 - \text{cost}_{AQ}$</td>
<td>-1</td>
</tr>
</tbody>
</table>

Policy Gradient

$$\nabla J(\theta) = \nabla \log p_{RL}(a)[r - b]$$
Setting 2: Reinforcement Learning

Memory Network

Policy Gradient

Baseline

\[ \nabla J(\theta) = \nabla \log p_{RL}(a)[r - b] \]
Q&A