Neural Dialogue Generation

Jiwei Li
Computer Science Department
Stanford University
Collaborators

Michel Galley
Microsoft Research

Jianfeng Gao
Microsoft Research

Dan Jurafsky
Stanford

Bill Dolan
Microsoft Research

Chris Brockett
Microsoft Research

Alan Ritter
Ohio State University
Learn to Converse

Where is *The Hobbit* playing in Mountain View?

The Hobbit is playing at the Century 16 Theater.

When is it playing there?

It's playing at 2pm, 5pm, and 8pm.

OK. I'd like 1 adult and 2 children for the first show. How much would that cost?

I can't deal with it *atm*...

Because of your game?

Yeah, I'm on my way now.

You're going now? Ok good luck!
Learn to Converse

ChatBot

Where is *The Hobbit* playing in Mountain View?

*The Hobbit* is playing at the Century 16 Theater.

When is it playing there?

It's playing at 2pm, 5pm, and 8pm.

OK. I'd like 1 adult and 2 children for the first show. How much would that cost?

I can't deal with it atm...

Because of your game?

Yeah, I'm on my way now

You're going now? Ok good luck!
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?

Slid Borrowed From Bill MacCartney
Why is building a chatbot hard?
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.
Background
Background

Template/Rule based systems (Levin et al., 2000; Young et al., 2010; Walker et al., 2003; Pieraccini et al., 2009; Wang et al., 2011)
Background

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)
Neural Generation Models for MT
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

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

Target: French
Source: English

What's your name?
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

Loss = − log p(target|source)

Source : Input Messages
Target : Responses

Encoding

Decoding

how are you ?

I’m fine .

eos I’m fine .

EOS
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

Long Short Term Memory

Encoding

\[
\begin{align*}
  i_t &= \sigma(W_i \cdot [h_{t-1}, e_t]) \\
  f_t &= \sigma(W_f \cdot [h_{t-1}, e_t]) \\
  o_t &= \sigma(W_o \cdot [h_{t-1}, e_t]) \\
  l_t &= \tanh(W_l \cdot [h_{t-1}, e_t]) \\
  c_t &= f_t \cdot c_{t-1} + i_t \cdot l_t \\
  h_t^o &= o_t \cdot \tanh(c_t)
\end{align*}
\]
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)
Neural Generation Models as a Backbone

I'm fine.
Outline

1. Mutual Information for Response Generation.
   The “I don’t know” problem
2. Speaker Consistency
3. Multi-context Response Generation
4. Reinforcement learning for Response Generation
Mutual Information for Response Generation.
Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL, 2016)
Mutual Information for Response Generation.

Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL, 2016)

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015)
Mutual Information for Response Generation. Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL, 2016)

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015; )
Mutual Information for Response Generation.
Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL, 2016)

“I don’t know“ problem (Sordoni et al., 2015; Serban et al., 2015; )
Mutual Information for Response Generation.
Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL, 2016)

“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.

30% percent of all generated responses
Mutual Information for Response Generation.

def ChatBot(input_string):
    if string[len(input_string)-1]=='?':
        return "I don't know";
    else:
        return "I don't know what you are talking about";
Mutual Information for Response Generation.

Solution #1: Adding Rules
Mutual Information for Response Generation.

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 ! ! !
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 haven’t the faintest idea

I don’t have the foggiest idea what you are talking about.

I don’t have the lightest idea what you are talking about.

How should I know?
Mutual Information for Response Generation.

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 ?

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

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

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

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

“I don’t know”\hspace{1cm}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.

\[ \mathbb{P}(\text{target} \mid \text{source}) \]

What one asks \[\rightarrow\] “I don’t know”

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

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.

\[ \hat{T} = \arg \max_T \{ \log \frac{p(S, T)}{p(S)p(T)} \} \]

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

Bayesian Rule
Mutual Information for Response Generation.

\[
\hat{T} = \arg \max_T \left\{ \log \frac{p(S, T)}{p(S)p(T)} \right\}
\]

Bayesian Rule

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

Standard Seq2Seq model
Mutual Information for Response Generation.

\[
\hat{T} = \arg \max_T \left\{ \log \frac{p(S, T)}{p(S)p(T)} \right\}
\]

Bayesian Rule

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

Anti-language Model
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.

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

Bayesian Rule

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

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

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

Bayesian Rule

\[ T = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T') \right\} \]

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

Solution 1

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

Solution 2

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

Datasets: Twitter conversational Dataset (23M pairs)
          Opensubtitle movie scripts dataset (80M pairs)
Datasets and Evaluations

Datasets: Twitter conversational Dataset (23M pairs)  
Opensubtitle movie scripts dataset (80M pairs)

Evaluations: BLEU (Papineni et al., 2003)  
#Distinct tokens  
Human Evaluation  
(1000 samples, each output is evaluated by 7 judges)
Datasets and Evaluations

**BLEU**

- **Ritter’s MT system**
  - +22.5%

- **Seq2Seq**
  - +35.0%

- **p(t|s)-p(t)**
  - +12.7%

- **p(t|s)+p(s|t)**
  - +26.4%
  - +51.3%
Datasets and Evaluations

BLEU on Twitter Dataset

![Bar chart showing BLEU scores for different models.](chart.png)
Datasets and Evaluations

# Distinct Tokens in generated targets (divided by total #) on Opensubtitle dataset

+385%  +122%
## Human Evaluation

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Gain</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMT (Ritter et al., 2011)</td>
<td>0.29</td>
<td>[0.25, 0.32]</td>
</tr>
<tr>
<td>SMT+neural reranking</td>
<td>0.28</td>
<td>[0.25, 0.32]</td>
</tr>
<tr>
<td>SEQ2SEQ (baseline)</td>
<td>0.11</td>
<td>[0.07, 0.14]</td>
</tr>
<tr>
<td>SEQ2SEQ (greedy)</td>
<td>0.08</td>
<td>[0.04, 0.11]</td>
</tr>
</tbody>
</table>

## Human Evaluation

<table>
<thead>
<tr>
<th>Comparator</th>
<th>Gain</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMT (Ritter et al., 2011)</td>
<td>0.29</td>
<td>[0.25, 0.32]</td>
</tr>
<tr>
<td>SMT+neural reranking</td>
<td>0.28</td>
<td>[0.25, 0.32]</td>
</tr>
<tr>
<td><strong>SEQ2SEQ (baseline)</strong></td>
<td>0.11</td>
<td>[0.07, 0.14]</td>
</tr>
<tr>
<td><strong>SEQ2SEQ (greedy)</strong></td>
<td>0.08</td>
<td>[0.04, 0.11]</td>
</tr>
</tbody>
</table>
## Sampled Results

**Input: What are you doing?**

| Sampled Seq2Seq $p(t|s)$         | Mutual Information          |
|---------------------------------|-----------------------------|
| **Input:** What are you doing?  |                              |
| -0.86 I don’t know.             | 1. I’ve been looking for you.|
| -1.09 Get out of here.          | 4. I told you to shut up.    |
| -1.03 I don’t know!             | 2. I want to talk to you.    |
| -1.06 Nothing.                  | 5. Get out of here.          |
| -1.09 Get out of the way.       | 3. Just making sure you’re OK.|
| -1.10 I’m talking to you.       | 6. I’m looking for a doctor. |

**Input: What is your name?**

| Sampled Seq2Seq $p(t|s)$         | Mutual Information          |
|---------------------------------|-----------------------------|
| **Input:** What is your name?   |                              |
| -0.91 I don’t know.             | 1. Blue!                    |
| -0.92 I don’t know!             | 4. Daniel.                  |
| -0.92 I don’t know, sir.        | 2. Peter.                   |
| -0.97 Oh, my god!               | 5. My name is John.         |
| -1.55 My name is Robert.        | 3. Tyler.                   |
| -1.58 My name is John.          | 6. My name is Robert.       |

**Input: How old are you?**

| Sampled Seq2Seq $p(t|s)$         | Mutual Information          |
|---------------------------------|-----------------------------|
| **Input:** How old are you?     |                              |
| -0.79 I don’t know.             | 1. Twenty-eight.            |
| -1.06 I’m fine.                 | 4. Five.                    |
| -1.17 I’m all right.            | 2. Twenty-four.             |
| -1.17 I’m not sure.             | 5. 15.                     |
| -1.64 Twenty-five.              | 3. Long.                    |
| -1.66 Five.                     | 6. Eight.                   |

Standard Seq2Seq $p(t|s)$
Outlines

1. Mutual Information for Response Generation.
2. Speaker Consistency
3. Multi-context Response Generation
4. Reinforcement learning for Response Generation
Speaker Consistency
Li et al., 2016. A Persona-Based Neural Conversation Model,
Speaker Consistency

Mutual Information

\[ \hat{T} = \arg \max_T \left\{ \log \frac{p(S, T)}{p(S)p(T)} \right\} \]

MLE

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

How old are you?

I’m 8.
Speaker Consistency

How old are you?

I’m 8.

What’s your age?

18
Speaker Consistency

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.
Speaker Consistency

Where do you live now?

I live in Los Angeles.

In which city do you live now?

I live in Paris.

In which country do you live now?

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?
Speaker Consistency

When were you born?

In 1942.
Speaker Consistency

When were you born?

In 1942.

When was your mother born?

In 1966.
Speaker Consistency
Speaker Consistency

1. Background facts, user profiles.
Speaker Consistency

1. Background facts, user profiles.
2. Language behavior
Speaker Consistency

1. Background facts, user profiles.
2. Language behavior
3. Interaction style
Speaker Consistency

1. Background facts, user profiles.
2. Language behavior
3. Interaction style

How to model speakers??
How to represent users

Word embeddings (50k)
How to represent users

Word embeddings (50k)

- uk
- london
- sydney
- great
- good
- monday
- live
- okay
- stay

Persona embeddings (70k)

- Bob
Persona seq2seq model

Encoding

Decoding

where  do  you  live  EOS
Persona seq2seq model

Encoding

where

do

you

live

Decoding

EOS

Bob

Persona embeddings (70k)
Persona seq2seq model

Encoding

where  do  you  live

Decoding

EOS

in

Bob

Persona embeddings (70k)
Persona seq2seq model

Encoding

where do you live

Decoding

EOS

in Bob

uk

Bob

Persona embeddings (70k)
Persona seq2seq model

Encoding

where  do   you  live

Decoding

in  uk  .

Bob

EOS

Persona embeddings (70k)

Bob
Persona seq2seq model

Encoding
- where
- do
- you
- live
- EOS

Decoding
- in
- uk
- .
- Bob
- in
- Bob
- uk
- Bob
- .
- Bob

Bob

Word embeddings (50k)
- great
- good
- sydney
- london
- monady
- tuesday
- live
- okay
- stay

Persona embeddings (70k)
- [Image]
Interaction Seq2Seq model

Encoding

where  do  you  live
Interaction Seq2Seq model

Encoding

where  do  you  live

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

\[
\text{where} \quad \text{do} \quad \text{you} \quad \text{live} \quad \text{EOS} \\
\tanh(W*)
\]
Interaction Seq2Seq model

Encoding:
- where
- do
- you
- live

Decoding:
- in
- uk

$tanh(W^* \begin{bmatrix} \cdot \cdot \cdot \\ \cdot \cdot \cdot \end{bmatrix})$
Interaction Seq2Seq model

where do you live

Encoding

Decoding

in

uk

in

uk

EOS

[Diagram showing the interaction between encoding and decoding processes, with the words "where", "do", "you", "live", "in", and "uk" connected through arrows.]

[Images of a smiling woman and a smiling man.]
Datasets and Evaluations

Conversation from Twitter 28M turns
74,003 Users minimum of 60 conversational turns

Perplexity
BLEU (4,000 single reference)
Human Evaluation
## Quantitative Results

<table>
<thead>
<tr>
<th></th>
<th>Seq2Seq</th>
<th>Speaker Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perplexity</strong></td>
<td>47.2</td>
<td>42.2 (-10.6%)</td>
</tr>
<tr>
<td><strong>BLEU (without MMI)</strong></td>
<td>0.92</td>
<td>1.12 (+21.7%)</td>
</tr>
<tr>
<td><strong>BLEU (with MMI)</strong></td>
<td>1.41</td>
<td>1.66 (+11.7%)</td>
</tr>
</tbody>
</table>
Human Evaluation

Question Pairs
Human Evaluation

Question Pairs

What city do you live in?
What country do you live in?
Human Evaluation

Question Pairs

What city do you live in?
What country do you live in?

Are you vegan or vegetarian?
Do you eat beef?
Human Evaluation

Question Pairs

What city do you live in?
What country do you live in?

London/UK

London/US
Human Evaluation

Which Model produces more consistent answers?
Each item is given to 5 judges.
Ties are discarded

<table>
<thead>
<tr>
<th></th>
<th>Seq2Seq Model</th>
<th>Persona Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item1</td>
<td>+1</td>
<td></td>
</tr>
<tr>
<td>Item2</td>
<td></td>
<td>+1</td>
</tr>
</tbody>
</table>
# Human Evaluation

<table>
<thead>
<tr>
<th>Seq2Seq Model</th>
<th>Persona Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.84</td>
<td>1.33 (+34.7%)</td>
</tr>
</tbody>
</table>
Results (No cherry-picking)

<table>
<thead>
<tr>
<th>User1</th>
<th></th>
<th>User2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>message</strong></td>
<td>Where is your hometown?</td>
<td><strong>message</strong></td>
<td>Where is your hometown?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I'm from England.</td>
<td><strong>response</strong></td>
<td>I live in Indonesia.</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>Where are you from?</td>
<td><strong>message</strong></td>
<td>In which city do you live now?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I'm from England.</td>
<td><strong>response</strong></td>
<td>England, you?</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>In which city do you live now?</td>
<td><strong>message</strong></td>
<td>In which country do you live now?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I live in London.</td>
<td><strong>response</strong></td>
<td>I live in Jakarta.</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>In which country do you live now?</td>
<td><strong>message</strong></td>
<td>In which country do you live now?</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What is your major?</td>
<td><strong>message</strong></td>
<td>What is your major?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>Business. You?</td>
<td><strong>response</strong></td>
<td>Business, you?</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What did you study in college?</td>
<td><strong>message</strong></td>
<td>What did you study in college?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I did business studies.</td>
<td><strong>response</strong></td>
<td>Psychology, you?</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>How old are you?</td>
<td><strong>message</strong></td>
<td>How old are you?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I'm 18.</td>
<td><strong>response</strong></td>
<td>I'm 18.</td>
</tr>
<tr>
<td><strong>message</strong></td>
<td>What is your age?</td>
<td><strong>message</strong></td>
<td>What is your age?</td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I'm 18.</td>
<td><strong>response</strong></td>
<td>I'm 16.</td>
</tr>
</tbody>
</table>
### Results (No cherry-picking)

<table>
<thead>
<tr>
<th>User1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<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 <strong>England</strong>.</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 <strong>England</strong>.</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 <strong>London</strong>.</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 <strong>England</strong>.</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>Business. 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 business studies.</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>

<table>
<thead>
<tr>
<th>User2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>message</strong></td>
<td>Where is your hometown?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I live in <strong>Indonesia</strong>.</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 <strong>Jakarta</strong>.</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 <strong>Indonesia</strong>.</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>Business, 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>Psychology, you?</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 16.</td>
<td></td>
</tr>
</tbody>
</table>
Results (No cherry-picking)

<table>
<thead>
<tr>
<th>User1</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<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>You?</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>

<table>
<thead>
<tr>
<th>User2</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>message</strong></td>
<td>Where is your hometown?</td>
<td></td>
</tr>
<tr>
<td><strong>response</strong></td>
<td>I live in Indonesia.</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>England, you?</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 Jakarta.</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 Indonesia.</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>, 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>Psychology, you?</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 16.</td>
<td></td>
</tr>
</tbody>
</table>
Results (No cherry-picking)

User 1

<table>
<thead>
<tr>
<th>message</th>
<th>response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Where is your hometown?</td>
<td>I'm from England.</td>
</tr>
<tr>
<td>Where are you from?</td>
<td>I'm from England.</td>
</tr>
<tr>
<td>In which city do you live now?</td>
<td>I live in London.</td>
</tr>
<tr>
<td>What is your major?</td>
<td>Business. You?</td>
</tr>
<tr>
<td>What did you study in college?</td>
<td>I did business studies.</td>
</tr>
<tr>
<td>How old are you?</td>
<td>I'm 18.</td>
</tr>
</tbody>
</table>

User 2

<table>
<thead>
<tr>
<th>message</th>
<th>response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Where is your hometown?</td>
<td>I live in Indonesia.</td>
</tr>
<tr>
<td>Where are you from?</td>
<td>England, you?</td>
</tr>
<tr>
<td>In which city do you live now?</td>
<td>I live in Jakarta.</td>
</tr>
<tr>
<td>In which country do you live now?</td>
<td>I live in Indonesia.</td>
</tr>
<tr>
<td>What is your major?</td>
<td>Business, you?</td>
</tr>
<tr>
<td>What did you study in college?</td>
<td>Psychology, you?</td>
</tr>
<tr>
<td>How old are you?</td>
<td>I'm 18.</td>
</tr>
<tr>
<td>What is your age?</td>
<td>I'm 18.</td>
</tr>
</tbody>
</table>

response I'm 16.
### Results (No cherry-picking)

<table>
<thead>
<tr>
<th>User</th>
<th>Message</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>Where is your hometown?</td>
<td>I'm from England.</td>
</tr>
<tr>
<td></td>
<td>Where are you from?</td>
<td>I'm from England.</td>
</tr>
<tr>
<td></td>
<td>In which city do you live now?</td>
<td>I live in London.</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>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User 2</th>
<th>Message</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Where is your hometown?</td>
<td>I live in Indonesia.</td>
</tr>
<tr>
<td></td>
<td>England, you?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>In which city do you live now?</td>
<td>I live in Jakarta.</td>
</tr>
<tr>
<td></td>
<td>In which country do you live now?</td>
<td>I live in Indonesia.</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>Psychology, you?</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 16.</td>
</tr>
</tbody>
</table>
### Results (No cherry-picking)

<table>
<thead>
<tr>
<th><strong>User1</strong></th>
<th><strong>User2</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>message</strong> Where is your hometown?</td>
<td><strong>message</strong> Where is your hometown?</td>
</tr>
<tr>
<td><strong>response</strong> I’m from England.</td>
<td><strong>response</strong> I live in <strong>Indonesia</strong>.</td>
</tr>
<tr>
<td><strong>message</strong> Where are you from?</td>
<td><strong>message</strong> Where are you from?</td>
</tr>
<tr>
<td><strong>response</strong> I’m from England.</td>
<td><strong>response</strong> England, you?</td>
</tr>
<tr>
<td><strong>message</strong> In which city do you live now?</td>
<td><strong>message</strong> In which city do you live now?</td>
</tr>
<tr>
<td><strong>response</strong> I live in London.</td>
<td><strong>response</strong> I live in Jakarta.</td>
</tr>
<tr>
<td><strong>message</strong> In which country do you live now?</td>
<td><strong>message</strong> In which country do you live now?</td>
</tr>
<tr>
<td><strong>message</strong> What is your major?</td>
<td><strong>message</strong> What is your major?</td>
</tr>
<tr>
<td><strong>response</strong> Business. You?</td>
<td><strong>response</strong> Business, you?</td>
</tr>
<tr>
<td><strong>message</strong> What did you study in college?</td>
<td><strong>message</strong> What did you study in college?</td>
</tr>
<tr>
<td><strong>response</strong> I did business studies.</td>
<td><strong>response</strong> Psychology, you?</td>
</tr>
<tr>
<td><strong>message</strong> How old are you?</td>
<td><strong>message</strong> How old are you?</td>
</tr>
<tr>
<td><strong>response</strong> I’m 18.</td>
<td><strong>response</strong> I’m 18.</td>
</tr>
<tr>
<td><strong>message</strong> What is your age?</td>
<td><strong>message</strong> What is your age?</td>
</tr>
<tr>
<td><strong>response</strong> I’m 18.</td>
<td><strong>response</strong> I’m 16.</td>
</tr>
</tbody>
</table>
Persona seq2seq model

Encoding

where

do

you

live

Decoding

in

uk

Bob

Bob

Bob

Bob

EOS

Tradeoff
Outlines

1. Mutual Information for Response Generation.
2. Speaker Consistency
3. Multi-context Response Generation
4. Reinforcement learning for Response Generation
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 vocation.

Any particular plan?

????
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

**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?

Message: m

Response r

????
Multi-context Response Generation

Notations

Context c1
What's your plan for the upcoming summer?

Context c2
I am going to Hawaii for vocation.

... Any particular plan?

Message: m

Response r

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

I am going to Hawaii for vocation.
Multi-context Response Generation

What’s your plan for the upcoming summer?

I am going to Hawaii for vocation.
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.

$C = \sum_{i} O_i c_i$

Memory Network (Weston et al., 2014)
Multi-context Response Generation

Encoding

Decoding

where do you live

EOS

$C = \sum_i O_i c_i$

$c_1$ $c_2$ $c_K$
**Multi-context Response Generation**

![Diagram](image)

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

\[ C = \sum_i O_i c_i \]

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

- Encoding
- where
- do
- you
- live
- Decoding
- EOS
- C
Multi-context Response Generation

![Diagram showing encoding and decoding process with equations:]

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

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

Encoding

where

do

you

live

Decoding

m

in

C

uk

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

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

Encoding

where  do  you  live

Decoding

m  in  uk

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

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

where do you live

Encoding

Decoding

EOS

C

C

C

C

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

\[ C = \sum_i O_i c_i \]
Results on the Opensubtitle Dataset

Perplexity

# of context sentences

Perplexity vs. # of context sentences graph:

- Initial perplexity: 22.4
- Final perplexity: 17.1 (-23.6%)
Results on the Opensubtitle Dataset

![BLEU diagram](image)

# of context sentences
Outlines

1. Mutual Information for Response Generation.
2. Speaker Consistency
3. Multi-context Response Generation
4. Reinforcement learning for Response Generation
The goal of developing a ChatBot
The goal of developing a ChatBot
The goal of developing a ChatBot
The goal of developing a ChatBot

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

Data + Labels/Rewards
Supervised Learning

Data

You are a good boy.

You are a bad boy.

Labels
Supervised Learning

Data

Labels

Encoding

Decoding

What’s

your

name

?

eos

quel

est

votre

nom
What if labels or rewards is not immediate clear?
What if labels are not immediate clear?
What if labels are not immediate clear?
What if labels are not immediate clear?
Simulation

How old are you?
Simulation

How old are you?

I'm 16.
Simulation

How old are you?

16?

i'm 16.
Simulation

How old are you?

16?

I'm 16.

I don't know what you're talking about.
Simulation

How old are you?

16?

i'm 16.

i don't know what you're talking about

you don't know what you're saying
Simulation

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
Simulation

How old are you?

16?

I'm 16.

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

You don't know what you're saying.
Simulation

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

you don't know what you're saying
Simulation

How old are you?

16?

I'm 16.

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

You don't know what you're saying.

Bad Action
Reinforcement Learning

A set of environment states $S$ (current board for the Go game)
Reinforcement Learning

A set of environment states $S$

A set of actions to take a (where to place a stone)
Reinforcement Learning

A set of environment states $S$

A set of actions to take $a$ (where to place a stone)

Reward (capture the opponent’s stone, or win the entire game)
Reinforcement Learning

Goal: to learn which action to take given a specific state

By maximizing the overall reward function
Why can RL make the goal more achievable?

1. Allow us to design real world reward function
   1. Interesting vs not interesting; informative vs not informative
Why can RL make the goal more achievable?

1. Allow us to design real world reward function
   1. Interesting vs not interesting; informative vs not informative
   2. User feedback.
Can RL (to some extent) achieve this goal?

1. Allow us to design real world reward function
   1. Interesting vs not interesting; informative vs not informative
   2. User feedback
   3. Conversation Length ...
Can RL (to some extent) achieve this goal?

1. Allow us to design real world reward function
2. Look beyond two conversation turns (into the future)
Notations for Reinforcement Learning
Notations: Starting State

$r_{i-1}$

How old are you?

Encoding
Notations: Action

$r_{i-1}$

How old are you?

$r_i$

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

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

How old are you?

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

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

**How old are you?**

**i’m 16.**
Notations: **Reward**

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

- How old are you?
- i'm 16.
- tf-idf
Notations: **Reward**

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

- **tf-idf**
- **Informativeness**

How old are you?

i'm 16.
Notations: Reward

How old are you?

$i$'m 16.

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

**tf-idf**

# of turns it takes before generating dull responses
Notations: **Reward**

\[ R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \]

Mutual information (how old are you, I’m 16)
Objective function:

Goal: to learn which action to take given a specific state
Objective function:

Goal: to learn which action to take given a specific state

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

Based on the overall reward
Future Reward

$$\mathbb{E}_{r_i}[\hat{R}(r_i)] = \sum_{r_i} p(r_i|r_{i-1})[\hat{R}(r_i)]$$
\[ \mathbb{E}_{r_i}[\hat{R}(r_i)] = \sum_{r_i} p(r_i|r_{i-1})[\hat{R}(r_i)] \]
Action: Generating current response $r_i$

$$\mathbb{E}_{r_i}[\hat{R}(r_i)] = \sum_{r_i} p(r_i|r_{i-1})[\hat{R}(r_i)]$$

Approximation1: Sample a small list of candidates $A$
Action: Generating current response $r_i$

$$
\mathbb{E}_{r_i}[\hat{R}(r_i)] = \sum_{r_i} p(r_i|r_{i-1})[\hat{R}(r_i)]
$$

Approximation 1: Sample a small list of candidates $A$

$$
\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i|r_{i-1})}{\sum_{t \in A} p(r_i|r_{i-1})} [\hat{R}(r_i)]
$$

Normalization Part
Future Reward  \( \hat{R}(r_i) \)

\[
\hat{R}(r_i) = R(r_i) + \gamma \sum_{r_{i+1}} p(r_{i+1} | r_i) \hat{R}(r_{i+1})
\]

Immediate Reward
Future Reward \( \hat{R}(r_i) \)

\[
\hat{R}(r_i) = R(r_i) + \gamma \sum_{r_{i+1}} p(r_{i+1}|r_i) \hat{R}(r_{i+1})
\]

Immediate Reward

\[
\text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i)
\]
Future Reward

\[
\hat{R}(r_i) = R(r_i) + \gamma \sum_{r_{i+1}} p(r_{i+1}|r_i) \hat{R}(r_{i+1})
\]

Future Reward
Future Reward

\[ \hat{R}(r_i) = R(r_i) + \gamma \sum_{r_{i+1}} p(r_{i+1}|r_i) \hat{R}(r_{i+1}) \]

\[ \approx R(r_i) + \sum_{r_i \in A} \frac{p(r_{i+1}|r_i)}{\sum_{t \in A} p(r_{i+1}|r_i)} [\hat{R}(r_{i+1})] \]
Simulation
Simulation
Simulation

\[
\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \sum_{t \in A} \frac{p(r_i | r_{i-1})}{\sum_{t \in A} p(r_i | r_{i-1})} [\hat{R}(r_i)]
\]

\[
R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1} | r_i)}{\sum_{t \in A} p(r_{i+1} | r_i)} [R(r_{i+1})]
\]

A message from training set
Simulation

A message from training set

Decoding

\[
\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i|r_{i-1})}{\sum_{t \in A} p(r_i|r_{i-1})} [\hat{R}(r_i)]
\]

\[
R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1}|r_i)}{\sum_{t \in A} p(r_{i+1}|r_i)} [R(r_{i+1})]
\]
Simulation

A message from training set

$$\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i | r_{i-1})}{\sum_{t \in A} p(r_i | r_{i-1})} [\hat{R}(r_i)]$$

$$R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1} | r_i)}{\sum_{t \in A} p(r_{i+1} | r_i)} [R(r_{i+1})]$$
Simulation

$$\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i|r_{i-1})}{\sum_{t \in A} p(r_i|r_{i-1})} [\hat{R}(r_i)]$$

$$R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1}|r_i)}{\sum_{t \in A} p(r_{i+1}|r_i)} [R(r_{i+1})]$$
Simulation

\[
\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i | r_{i-1})}{\sum_{t \in A} p(r_i | r_{i-1})} [\hat{R}(r_i)]
\]

\[
R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1} | r_i)}{\sum_{t \in A} p(r_{i+1} | r_i)} [R(r_{i+1})]
\]
Simulation

\[
\text{Loss} = -\mathbb{E}_{r_i}[R(r_i)] \approx \sum_{r_i \in A} \frac{p(r_i|r_{i-1})}{\sum_{t \in A} p(r_i|r_{i-1})} [\hat{R}(r_i)]
\]

\[
R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in A} \frac{p(r_{i+1}|r_i)}{\sum_{t \in A} p(r_{i+1}|r_i)} [R(r_{i+1})]
\]
## 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 4.</td>
<td>I’m 4. why are you asking?</td>
</tr>
<tr>
<td></td>
<td></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 4.</td>
<td>I’m 4. 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>
</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 4.</td>
<td>I’m 4. 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>
</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 4.</td>
<td>I’m 4. 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

Keep your expectation low

How old are you?
Keep your expectation low

How old are you?

I'm 4, why are you asking?
Simulation

Keep your expectation low

How old are you?

I thought you were 12.

I'm 4, why are you asking?
Simulation

Keep your expectation low

How old are you?

I thought you were 12.

"I'm 4, why are you asking?"

What made you think so?
Simulation

Keep your expectation low

How old are you?

I'm 4. why are you asking?

I thought you were 12.

What made you think so?

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

Keep your expectation low

How old are you?

I thought you were 12.

I don’t know what you are talking about.

i’m 4, why are you asking?

What made you think so?

You don’t know what you are saying.
Simulation

Keep your expectation low

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 4, why are you asking?

What made you think so?

You don’t know what you are saying.
Simulation

Keep your expectation low

How old are you?

I thought you were 12.

I'm 4, why are you asking?

What made you think so?

You don't know what you are saying.

I don't know what you are talking about.

I don't know what you are talking about.
Conclusion

We apply neural generation methods to dialogue (chatbot) tasks
1. “I don’t know problem”
2. speaker consistency
3. leveraging multiple context information
4. preliminary attempt to make conversation keep flowing using RL
Q&A
Simulation

How old are you?

16?

I'm 16.

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

You don't know what you're saying.
Mutual Information for Response Generation.

Solution 1

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

Anti-language Model

1. Training \( P(T|S) \) and \( P(T) \)
2. Decoding

Ungrammatical Responses
Mutual Information for Response Generation.

Solution 1

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

\[ p(T) = \prod_{i=1}^{L_t} p(t_i|t_1, t_2, ..., t_{i-1}) \]
Mutual Information for Response Generation.

Solution 1

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

\[ p(T) = \prod_{i=1}^{L_t} p(t_i|t_1, t_2, \ldots, t_{i-1}) \]
Notations

Action: Generating current response $r_i$

State: LSTM hidden state obtained from history $r_1, r_2, ..., r_{i-1}$
Notations

Action: Generating current response $r_i$

State: LSTM hidden state obtained from history $r_1, r_2, ..., r_{i-1}$

transition function $p(r_i|r_1, r_2, ..., r_{i-1}, r_i) \approx p_\theta(r_i|r_{i-1}) \neq p_{\text{seq}}(r_{i+1}|r_i)$
Notations

Action: Generating current response $r_i$

State: LSTM hidden state obtained from history $r_1, r_2, \ldots, r_{i-1}$

transition function $p(r_i | r_1, r_2, \ldots, r_{i-1}, r_i) \approx p_\theta(r_i | r_{i-1})$

new State: LSTM hidden state obtained $r_1, r_2, \ldots, r_{i-1}, r_i$
Notations

Action: Generating current response \( r_i \)

State: LSTM hidden state obtained from history \( r_1, r_2, \ldots, r_{i-1} \)

transition function \( p(r_i | r_1, r_2, \ldots, r_{i-1}, r_i) \approx p_\theta(r_i | r_{i-1}) \)

new State: LSTM hidden state obtained \( r_1, r_2, \ldots, r_{i-1}, r_i \)

reward \( R(r_i) \)
Notations

Action: Generating current response $r_i$

State: LSTM hidden state obtained from history $r_1, r_2, \ldots, r_{i-1}$

transition function

$$p(r_i|r_1, r_2, \ldots, r_{i-1}, r_i) \approx p_\theta(r_i|r_{i-1})$$

new State: LSTM hidden state obtained $r_1, r_2, \ldots, r_{i-1}, r_i$

reward $R(r_i)$

$$R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i)$$
Notations

Action: Generating current response $r_i$

State: LSTM hidden state obtained from history $r_1, r_2, \ldots, r_{i-1}$

transition function $p(r_i | r_1, r_2, \ldots, r_{i-1}, r_i) \approx p_\theta(r_i | r_{i-1})$

new State: LSTM hidden state obtained $r_1, r_2, \ldots, r_{i-1}, r_i$

reward $R(r_i, s) = \text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i)$

$p_{\text{seq}}(t | s) \cdot p_{\text{seq}}(s | t)$
Notations

Action: Generating current response \( r_i \)

State S: LSTM hidden state obtained from history \( r_1, r_2, \ldots, r_{i-1} \)

Transition function \( p(r_i|r_1, r_2, \ldots, r_{i-1}) \approx p(r_i|r_{i-1}) \)

New State: LSTM hidden state obtained \( r_1, r_2, \ldots, r_{i-1}, r_i \)

Reward \( R(r_i, s) = 
\text{MMI}(r_i, r_{i-1}) \cdot \text{reward}(r_i) \)

Future reward \( \hat{R}(r_i, s) \)
Standard Seq2Seq2

Any particular plan?

Decoding