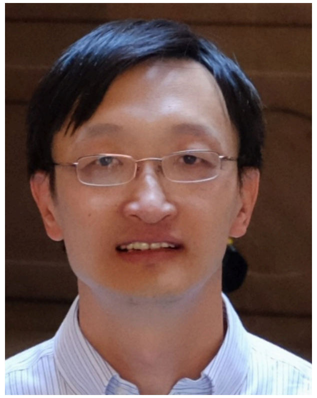
# Neural Dialogue Generation

Jiwei Li
Computer Science Department
Stanford University

# Collaborators

Michel Galley
Microsoft Research



Jianfeng Gao
Microsoft Research

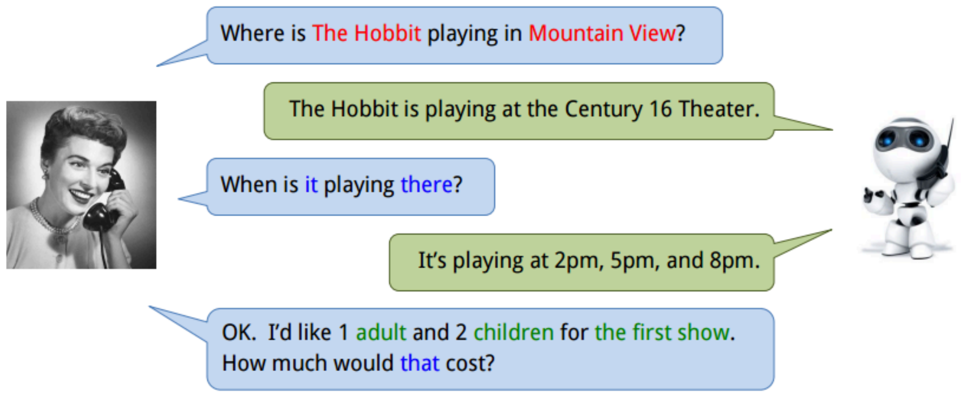
Dan Jurafsky
Stanford

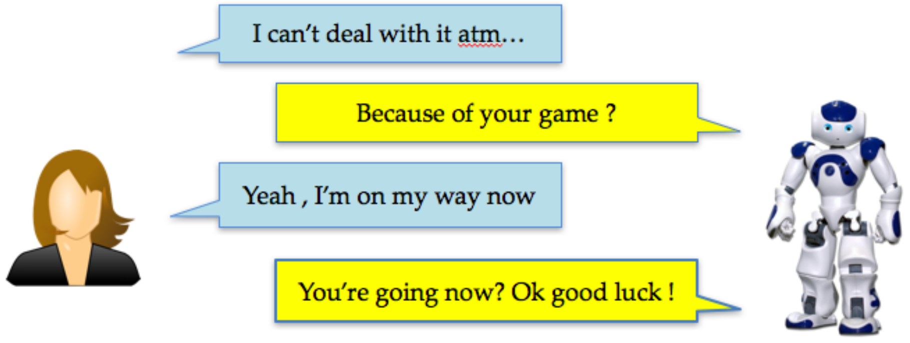
Bill Dolan
Microsoft Research

Chris Brockett
Microsoft Research

Alan Ritter
Ohio State University

# Learn to Converse



# 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 !

ChatBot


Hi. I'm Cortana.
Ask me a question!


Google
Now


amazon echo
amazon

# 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

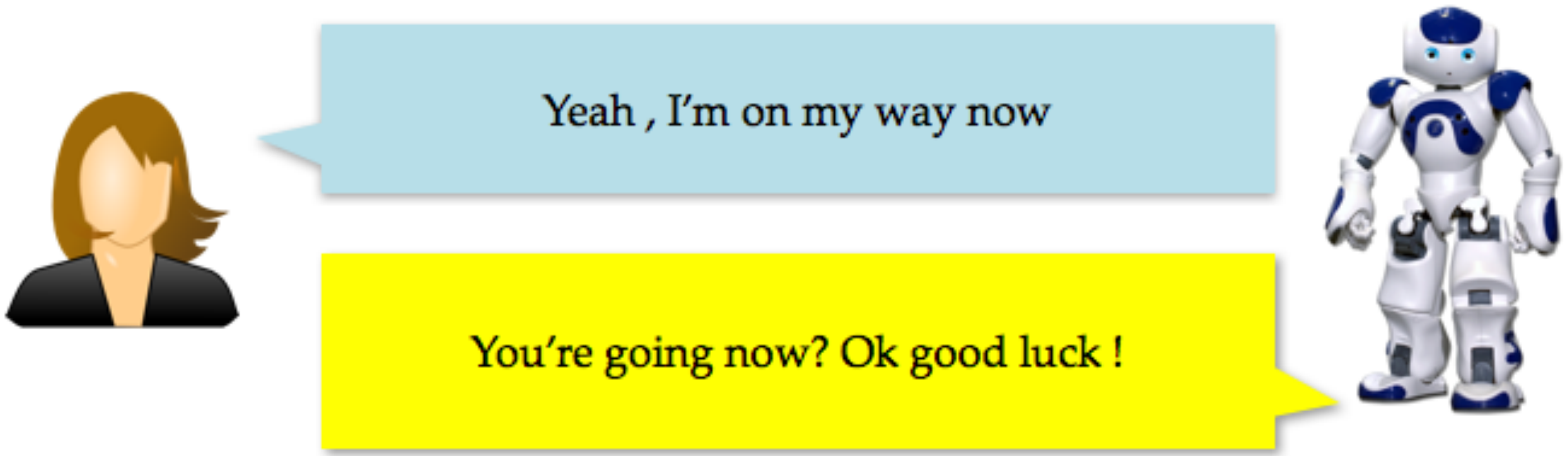
# Background

# 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

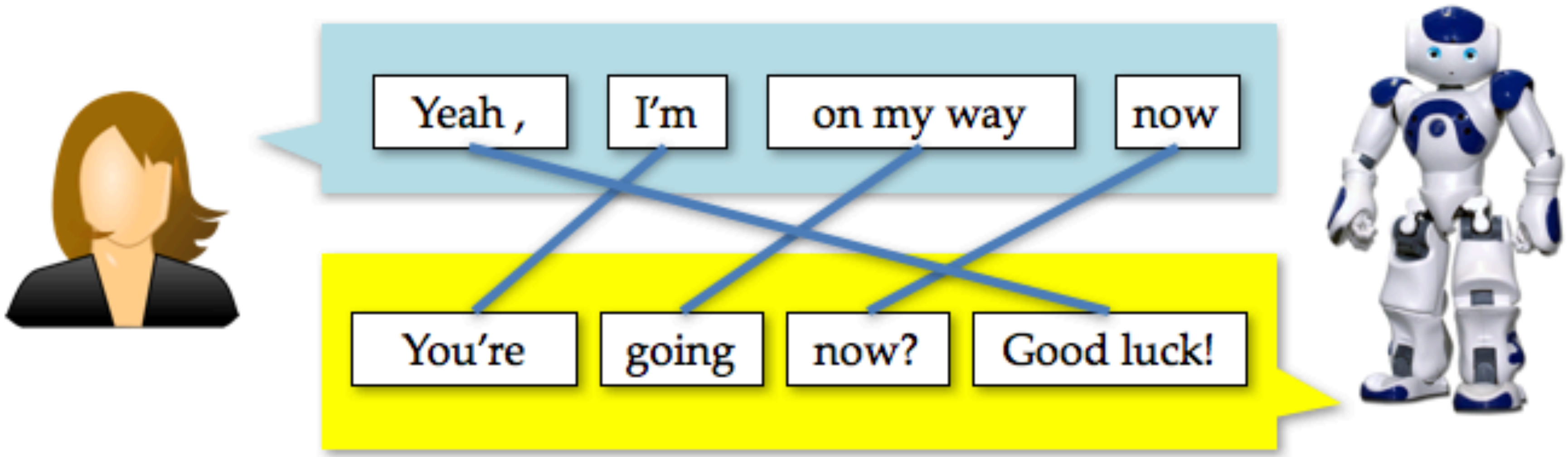
Response Generation as Statistical Machine Translation (Ritter et al., 2010)




Slide borrowed from Michel Galley

(Ritter et al., 2010)

# Response Generation as SMT




Exploit high-frequency patterns with phrase-based MT

"I am" → "you are"   "sick" → "get better"   "lovely!" → "thanks!"

Slide borrowed from Michel Galley

# Neural Generation Models for MT

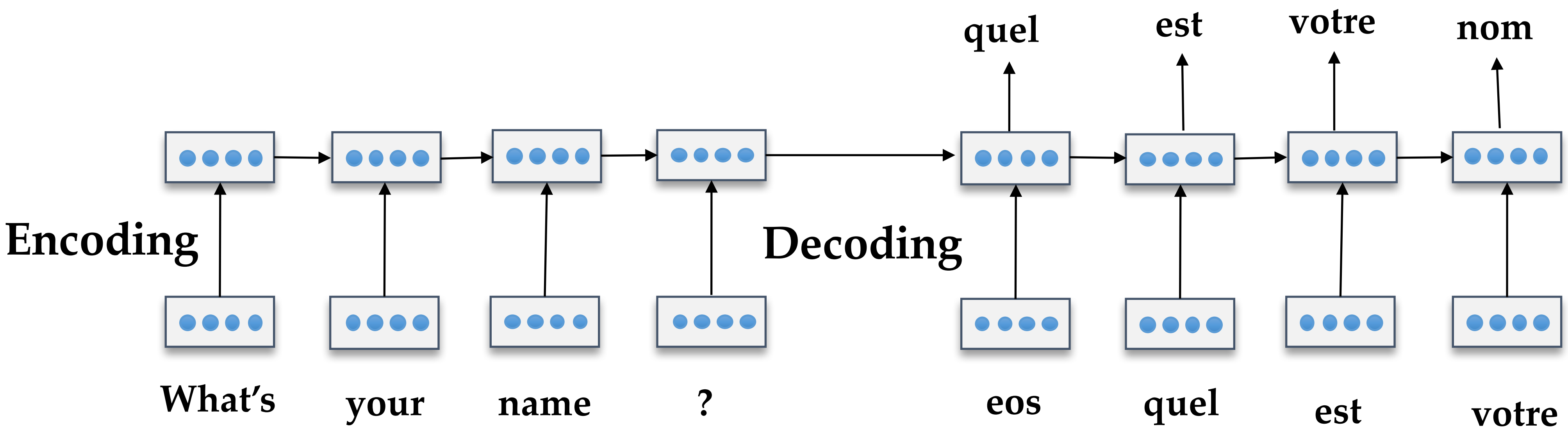
(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

# 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**

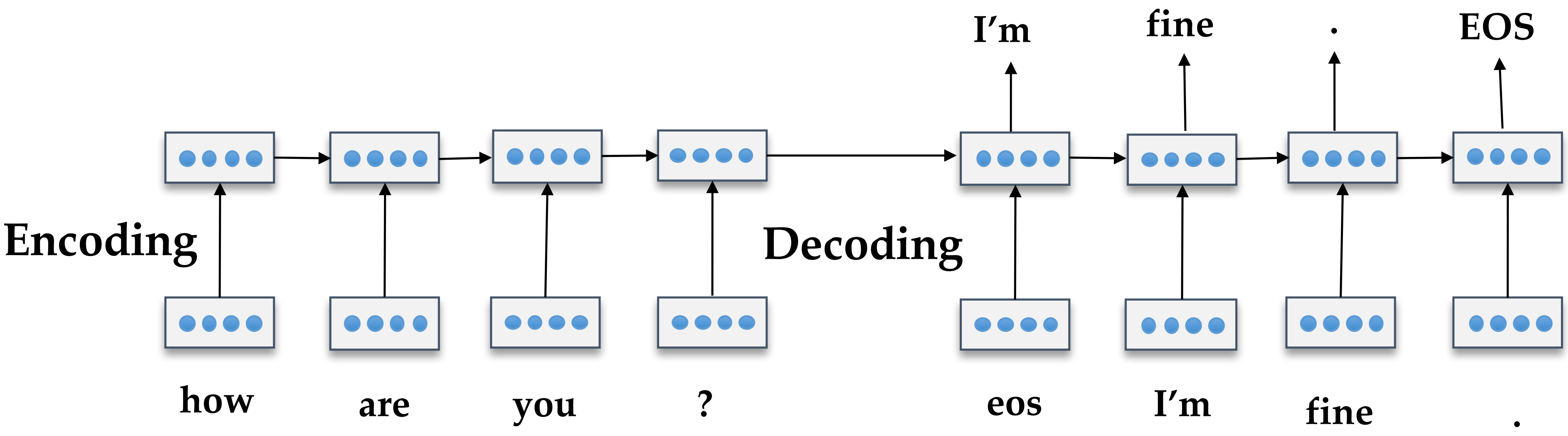


# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

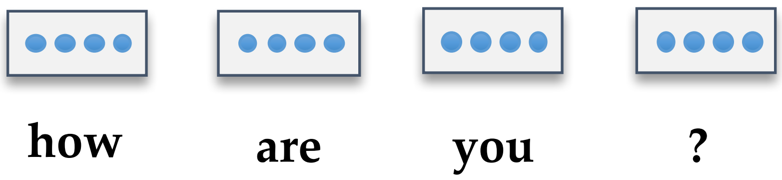
$$\text{Loss} = -\log p(\text{target}|\text{source})$$

**Source : Input Messages**
**Target : Responses**



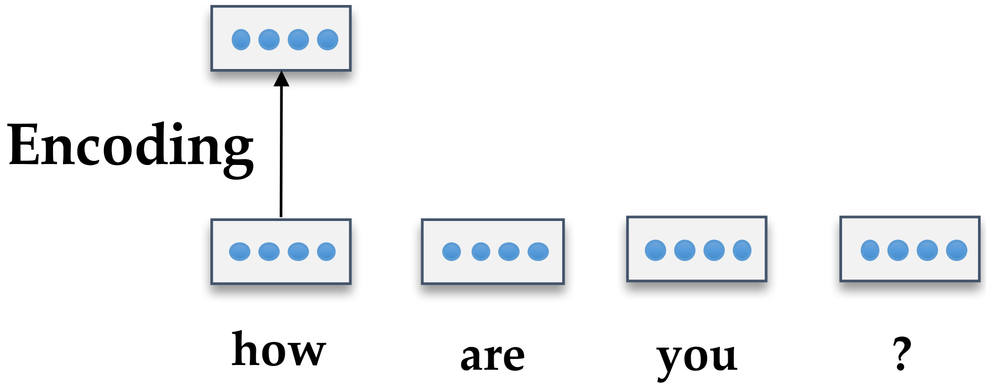
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



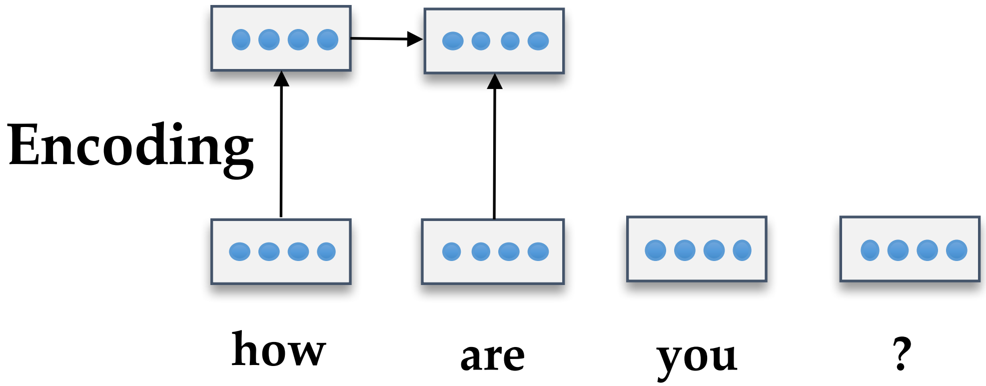
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



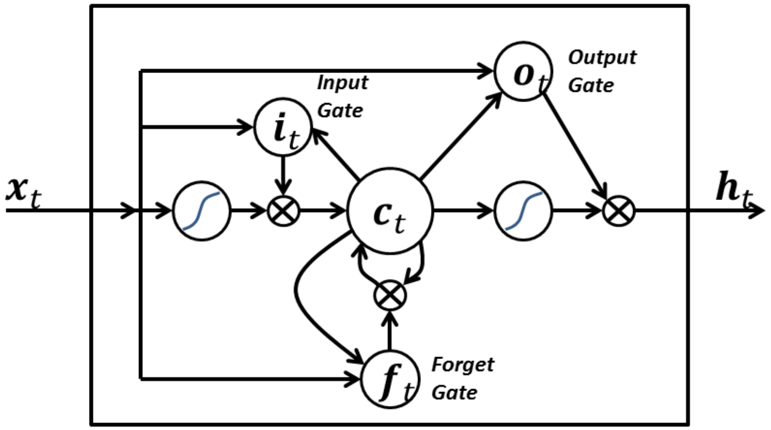
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



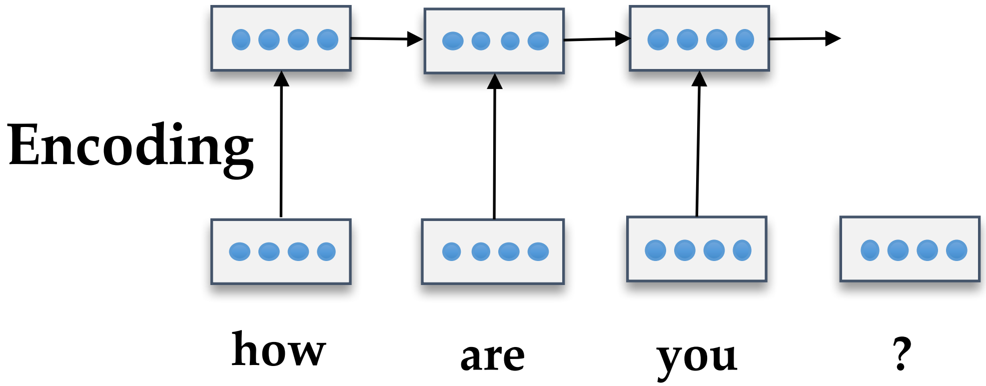

**Long Short Term Memory**

$$
\begin{aligned}
i_t &= \sigma(W_i \cdot [h_{t-1}, e_t]) \\
f_t &= \sigma(W_f \cdot [h_{t-1}, e_t]) \\
o_t &= \sigma(W_o \cdot [h_{t-1}, e_t]) \\
l_t &= \tanh(W_l \cdot [h_{t-1}, e_t]) \\
c_t &= f_t \cdot c_{t-1} + i_t \cdot l_t \\
h_t^s &= o_t \cdot \tanh(c_t)
\end{aligned}
$$



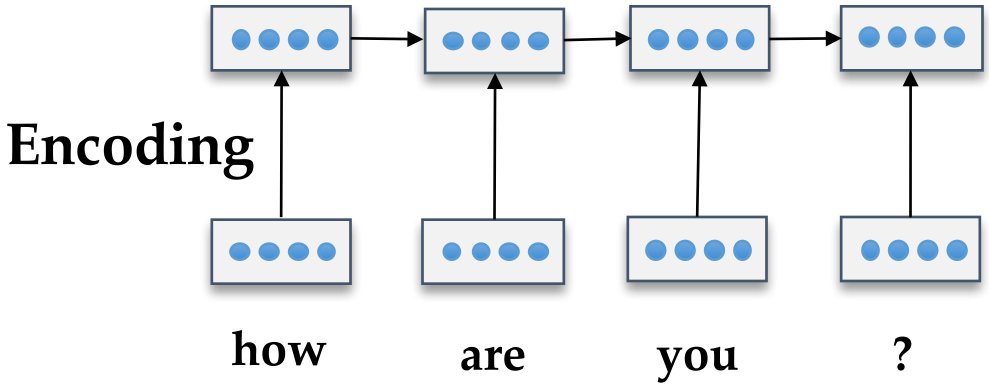
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



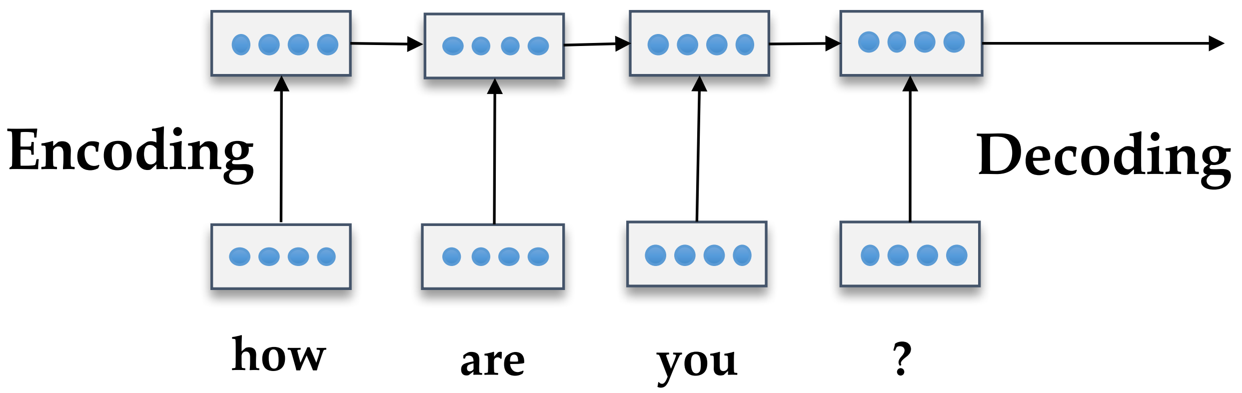
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



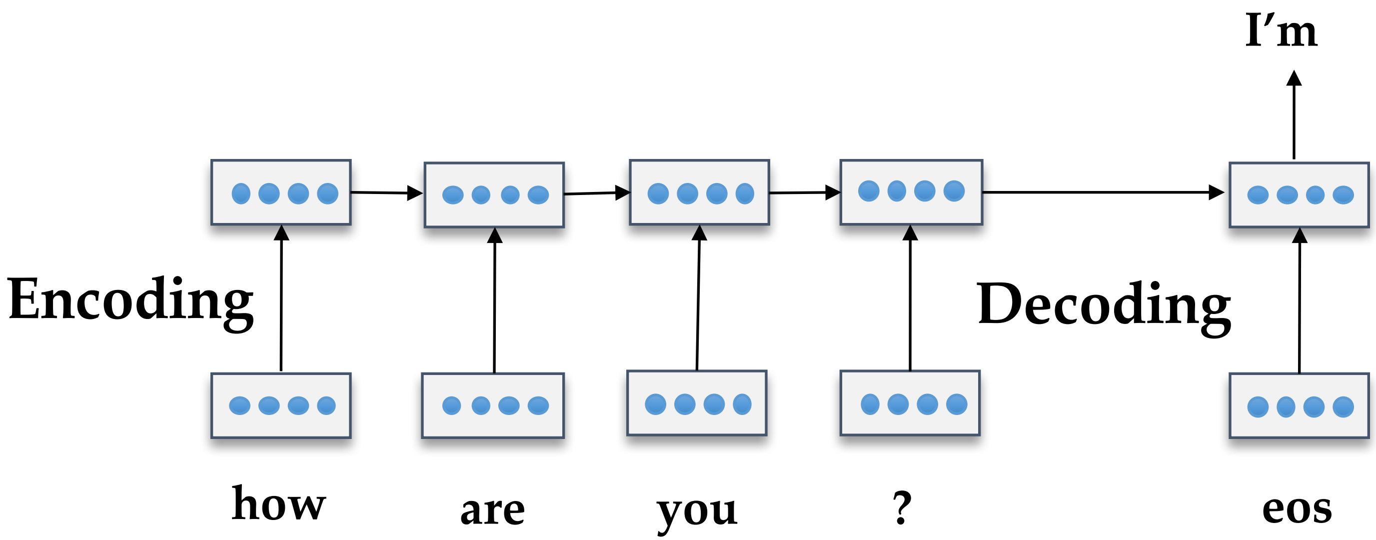
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



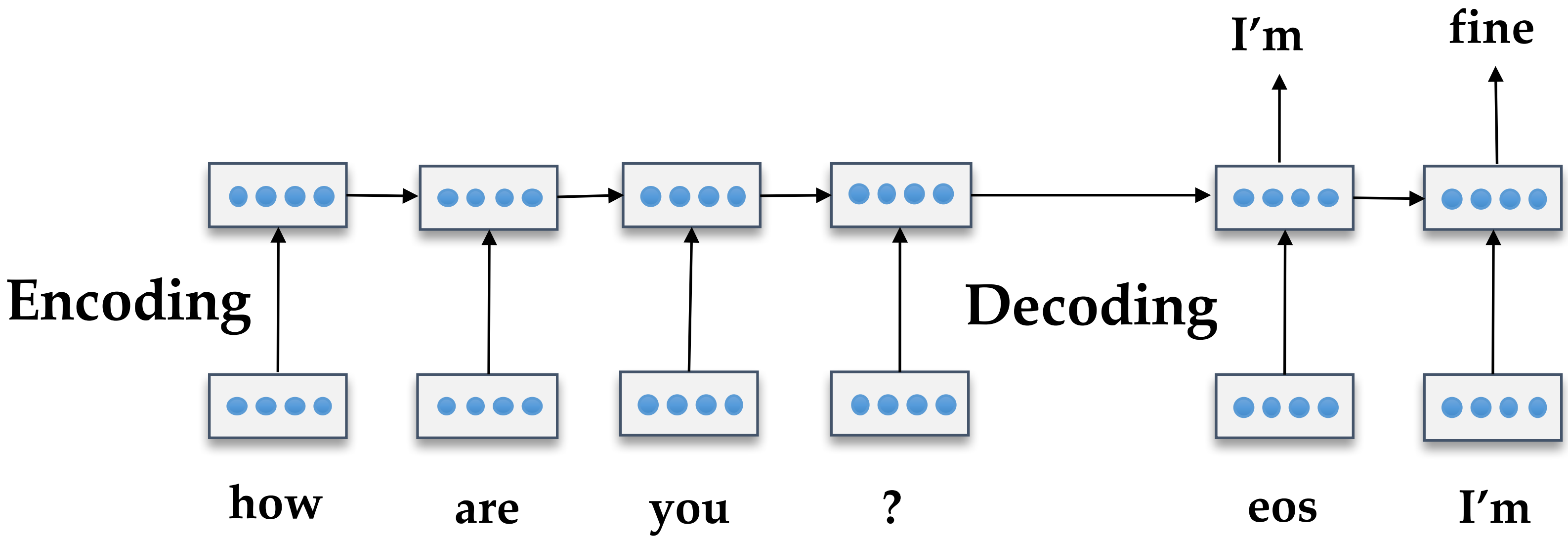
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



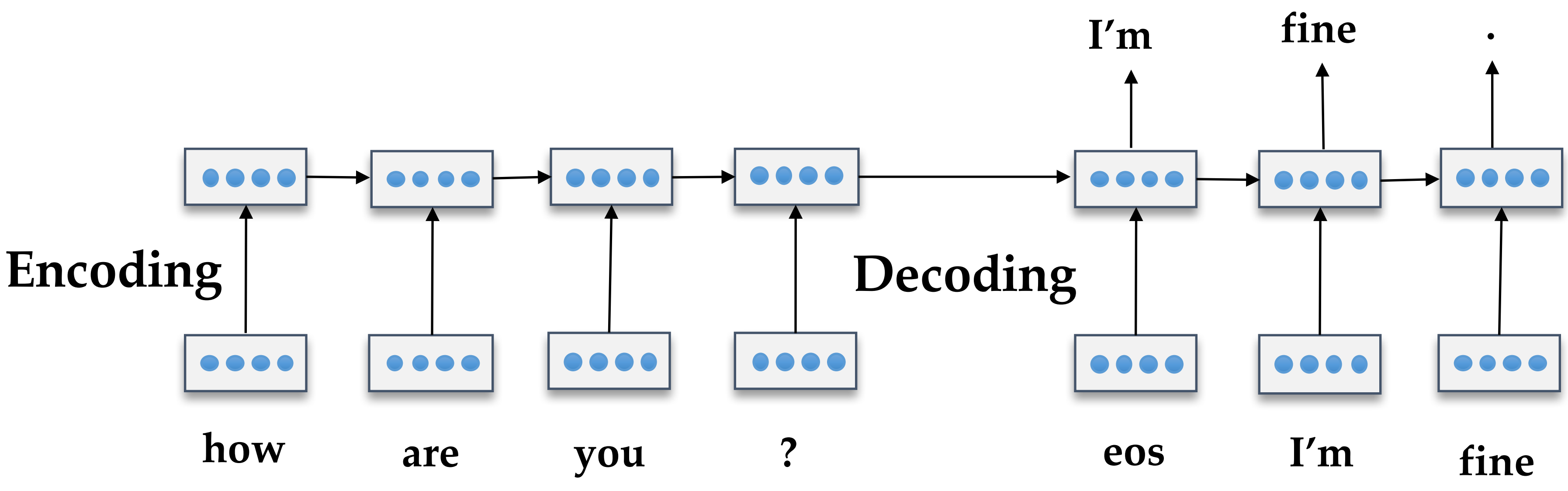
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



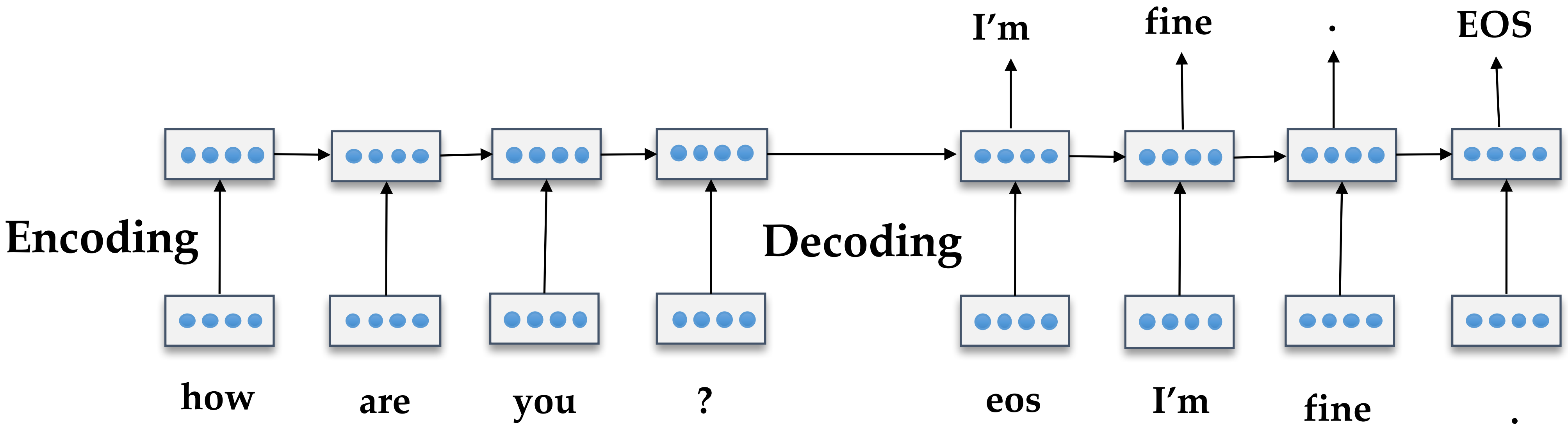
# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)

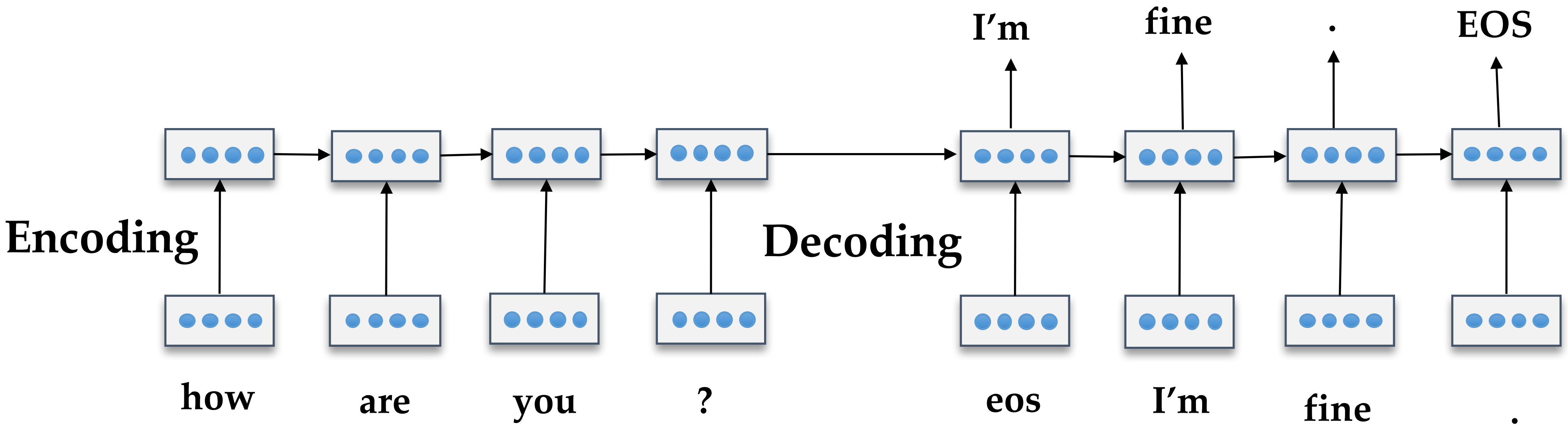


# Neural Generation Models

(Sutskever et al., 2014; Jean et al., 2014; Luong et al., 2015)



# Neural Generation Models as a Backbone



# 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.

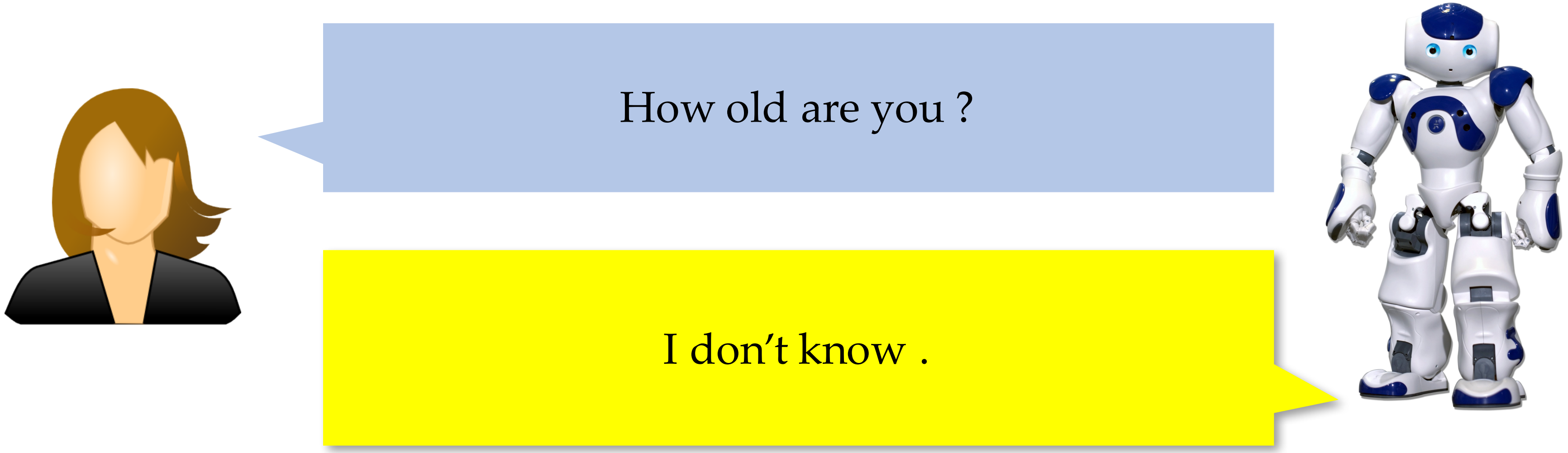
Li et al., A Diversity-Promoting Objective Function for Neural Conversation Models (to appear, NAACL,2016)

"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**)

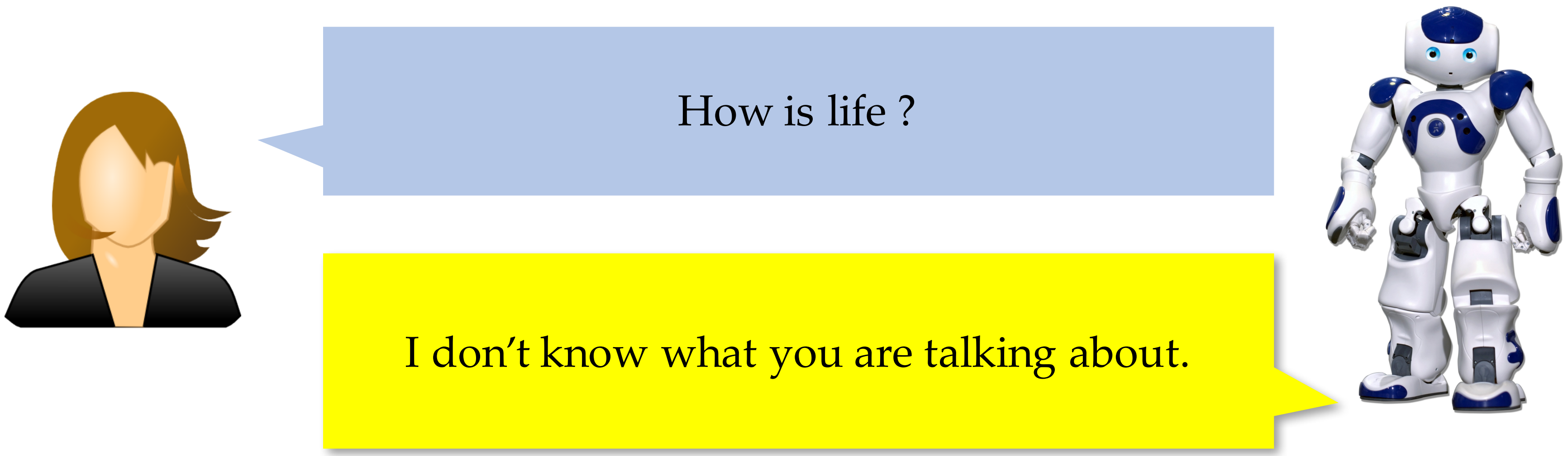
“I don’t know“ problem (Sordoni et al., 2015; Serban et al.,2015; )



# 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 ! ! !

# 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 the foggiest idea what you are talking about .

I don't have a clue.

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

I haven't the faintest idea

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 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 ?

**Rules don't work !!**

# Mutual Information for Response Generation.

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

# Mutual Information for Response Generation.

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

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

# 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.

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

What one asks $\longrightarrow$ "I don't know"

?

# Mutual Information for Response Generation.

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

What one asks → "I don't know"

?

"I don't know" → What one asks

?

# Mutual Information for Response Generation.

$$\text{source} \underset{p(\text{source}|\text{target})}{\overset{p(\text{target}|\text{source})}{\rightleftarrows}} \text{target}$$

# Mutual Information for Response Generation.

$$\text{source} \underset{p(\text{source}|\text{target})}{\overset{p(\text{target}|\text{source})}{\rightleftarrows}} \text{target}$$

**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.

$$\hat{T} = \arg\max_{T} \left\{ \log \frac{p(S,T)}{p(S)p(T)} \right\}$$

$$\downarrow \text{Bayesian Rule}$$

$$\hat{T} = \arg\max_{T} \left\{ \log p(T|S) - \boxed{\lambda} \log p(T) \right\}$$

# Mutual Information for Response Generation.

$$\hat{T} = \arg\max_{T} \left\{ \log \frac{p(S,T)}{p(S)p(T)} \right\}$$

$\downarrow$ Bayesian Rule

$$\hat{T} = \arg\max_{T} \left\{ \boxed{\log p(T|S)} - \lambda \log p(T) \right\}$$

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

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

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

$$T = \arg\max_{T} \left\{ (1-\lambda) \boxed{\log p(T|S)} + \lambda \log p(S|T) \right\}$$

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

# Mutual Information for Response Generation.

$$\hat{T} = \arg\max_{T} \left\{ \log p(T|S) - \lambda \log p(T) \right\}$$

$$\Big\downarrow \quad \text{Bayesian Rule}$$

$$T = \arg\max_{T} \left\{ (1-\lambda) \log p(T|S) + \lambda \boxed{\log p(S|T)} \right\}$$

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

# 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)

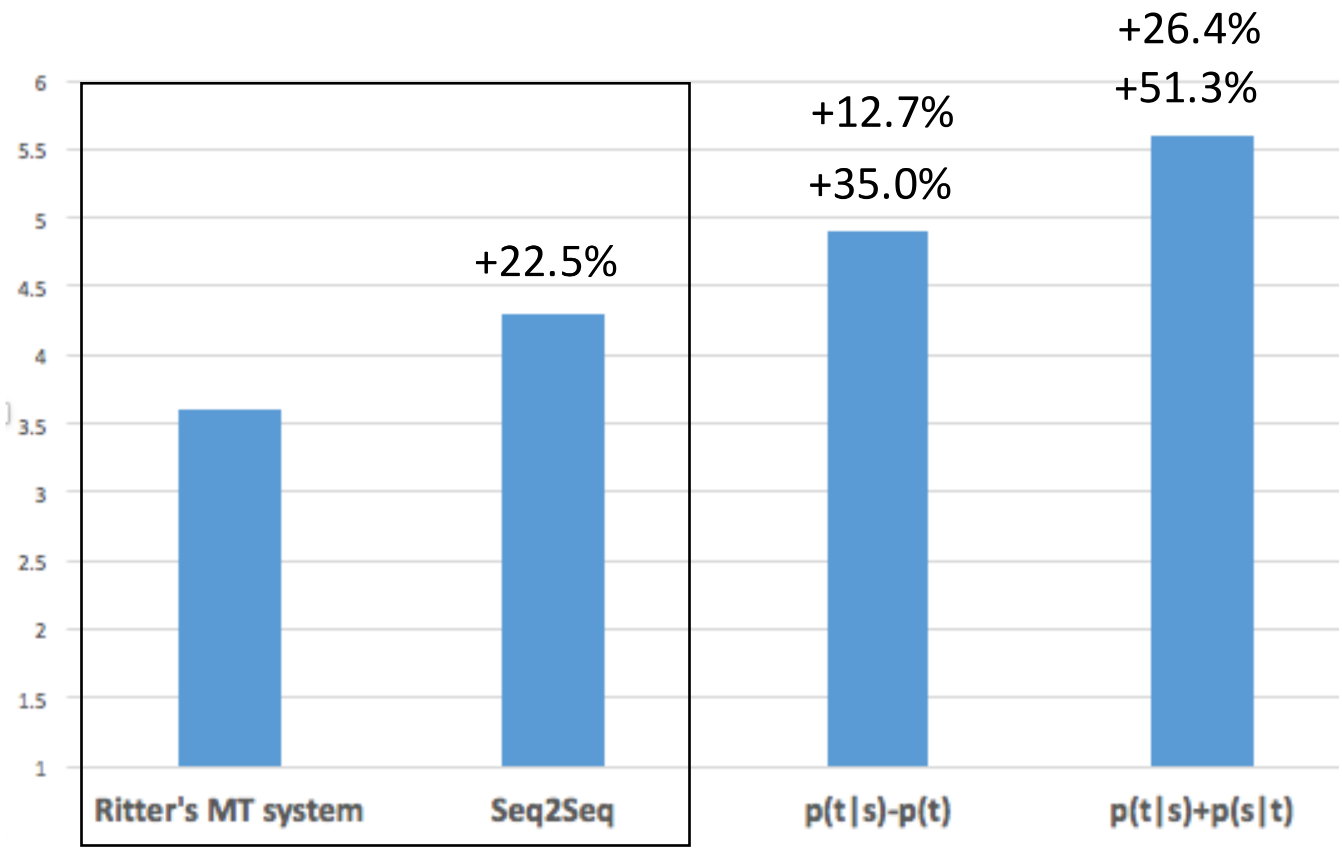
# Datasets and Evaluations

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)

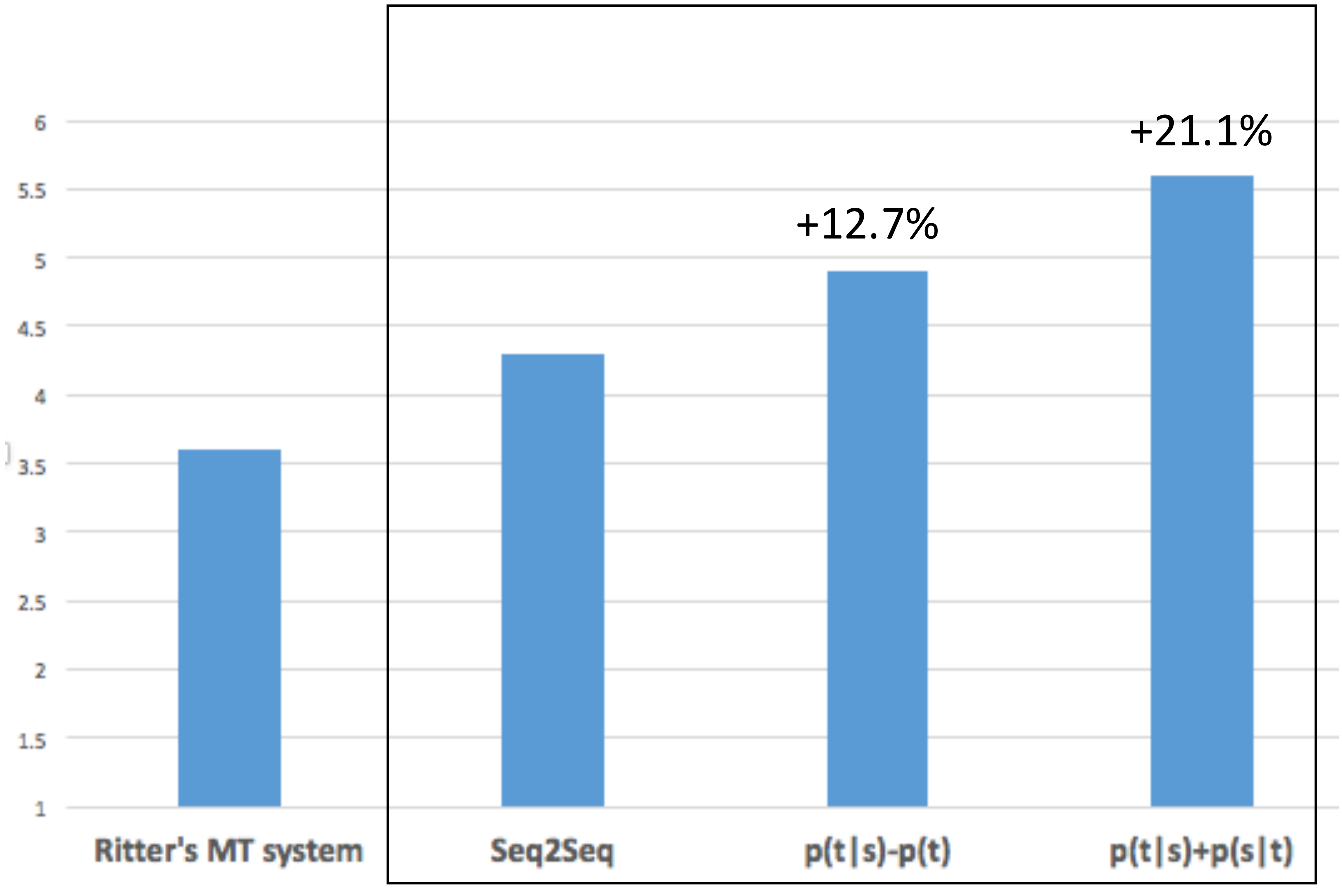
# Datasets and Evaluations

BLEU



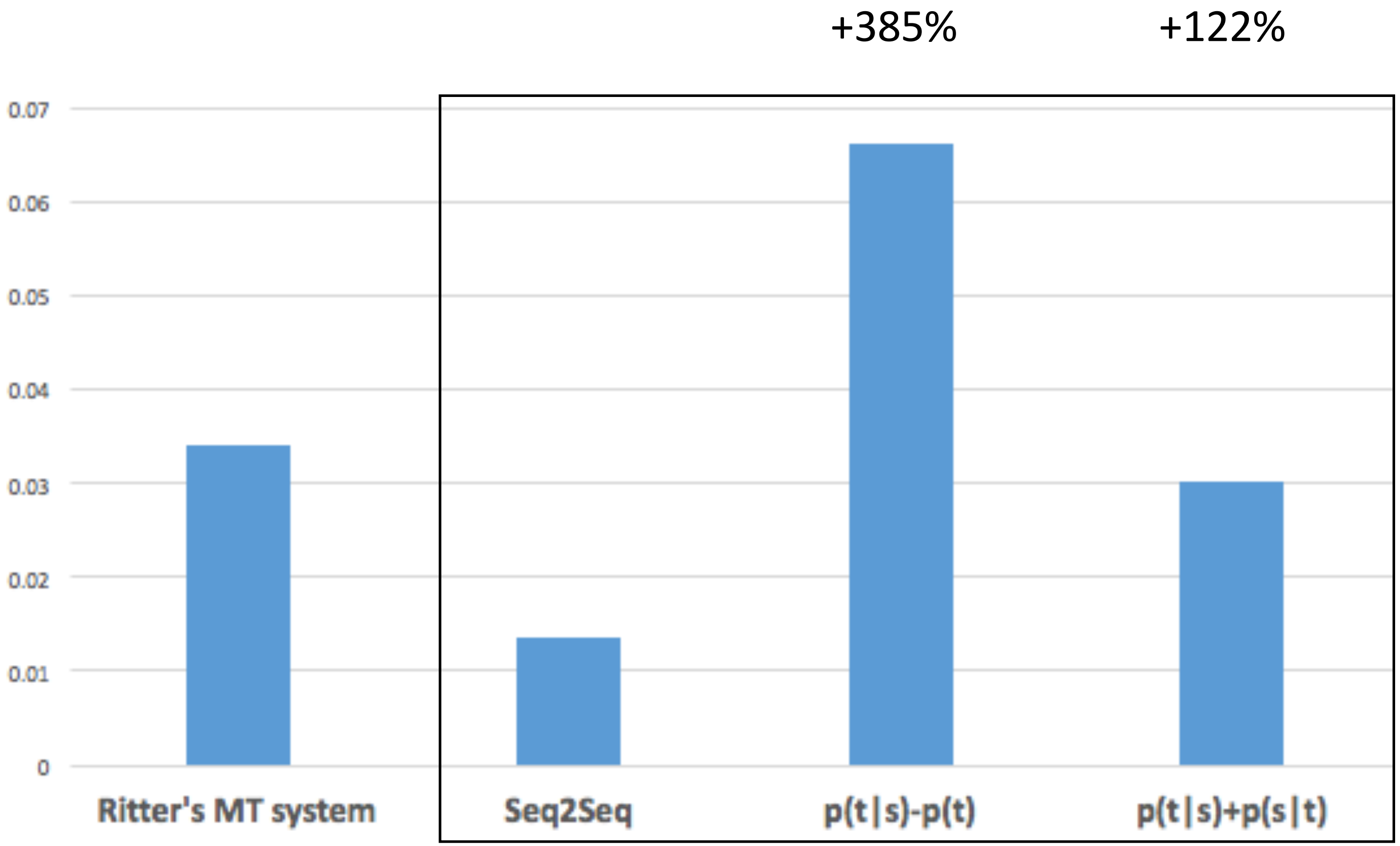
# Datasets and Evaluations

BLEU on Twitter Dataset



# Datasets and Evaluations

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



# Human Evaluation

| Comparator | Gain | 95% CI |
|---|---|---|
| SMT (Ritter et al., 2011) | 0.29 | [0.25, 0.32] |
| SMT+neural reranking | 0.28 | [0.25, 0.32] |
| SEQ2SEQ (baseline) | 0.11 | [0.07, 0.14] |
| SEQ2SEQ (greedy) | 0.08 | [0.04, 0.11] |

# Human Evaluation

| Comparator | Gain | 95% CI |
|---|---|---|
| SMT (Ritter et al., 2011) | 0.29 | [0.25, 0.32] |
| SMT+neural reranking | 0.28 | [0.25, 0.32] |
| SEQ2SEQ (baseline) | 0.11 | [0.07, 0.14] |
| SEQ2SEQ (greedy) | 0.08 | [0.04, 0.11] |

# Sampled Results

| Input: What are you doing? | | | |
|---|---|---|---|
| -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. |
| **Input**: what is your name? | | | |
| -0.91 | I don't know. | ... | |
| -0.92 | I don't know! | -1.55 | My name is Robert. |
| -0.92 | I don't know, sir. | -1.58 | My name is John. |
| -0.97 | Oh, my god! | -1.59 | My name's John. |
| **Input**: How old are you? | | | |
| -0.79 | I don't know. | ... | |
| -1.06 | I'm fine. | -1.64 | Twenty-five. |
| -1.17 | I'm all right. | -1.66 | Five. |
| -1.17 | I'm not sure. | -1.71 | Eight. |

Standard Seq2Seq p(t|s)

| Input: What are you doing? | |
|---|---|
| 1. I've been looking for you. | 4. I told you to shut up. |
| 2. I want to talk to you. | 5. Get out of here. |
| 3. Just making sure you're OK. | 6. I'm looking for a doctor. |
| **Input**: What is your name? | |
| 1. Blue! | 4. Daniel. |
| 2. Peter. | 5. My name is John. |
| 3. Tyler. | 6. My name is Robert. |
| **Input**: How old are you? | |
| 1. Twenty-eight. | 4. Five. |
| 2. Twenty-four. | 5. 15. |
| 3. Long. | 6. Eight. |

Mutual Information

# 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

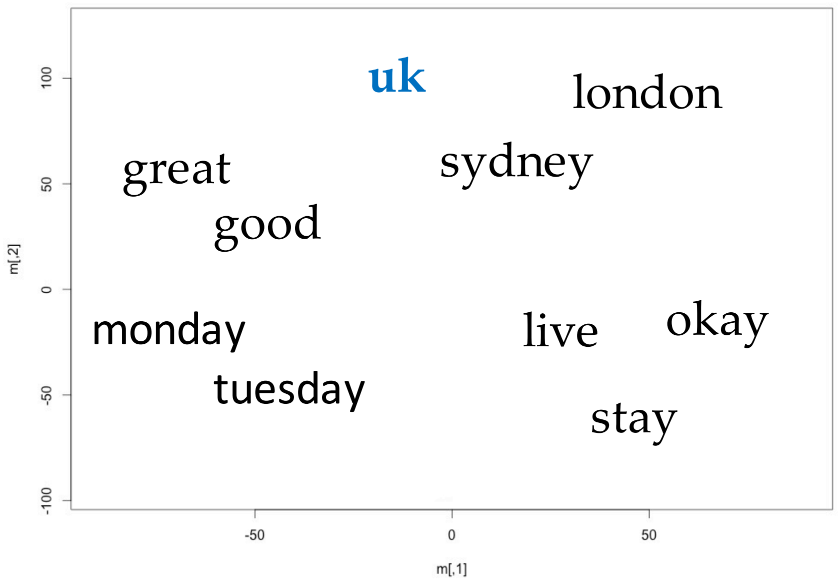
1. Background facts, user profiles.
2. Language behavior
3. Interaction style

# Speaker Consistency

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

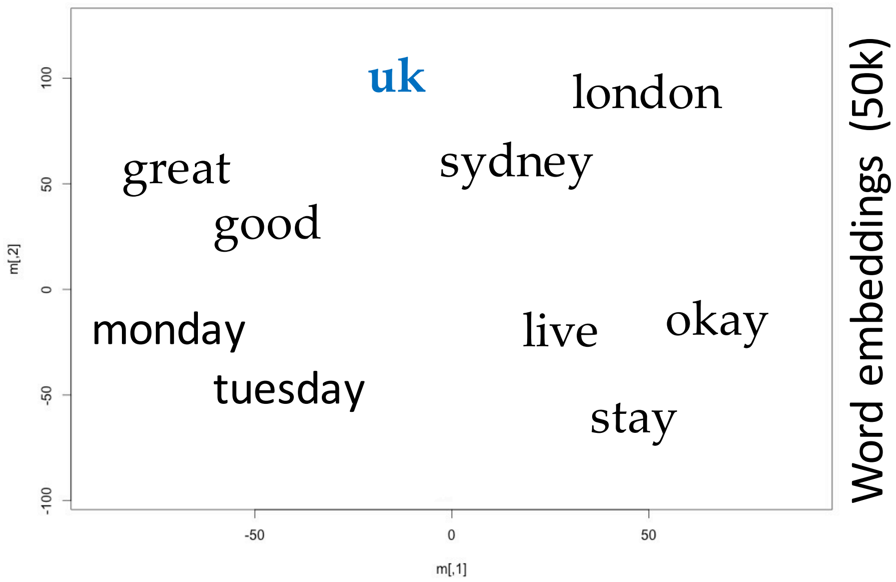
How to model speakers ??

# How to represent users

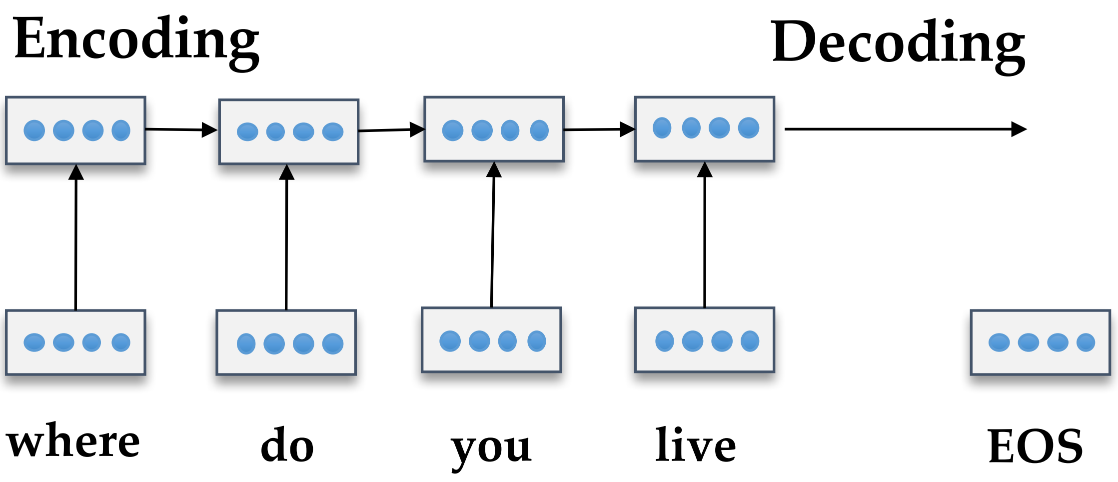



Word embeddings (50k)

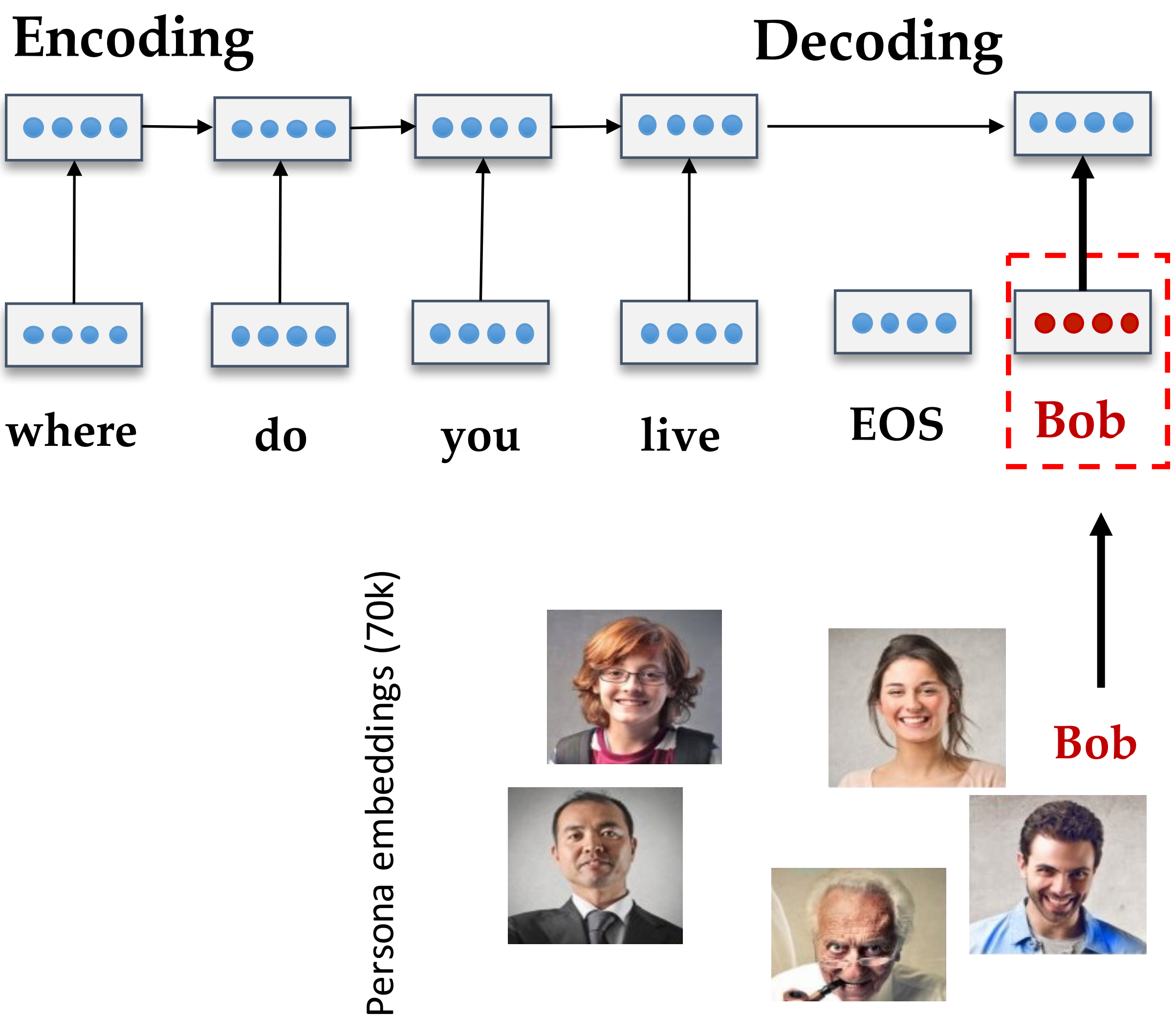
# How to represent users



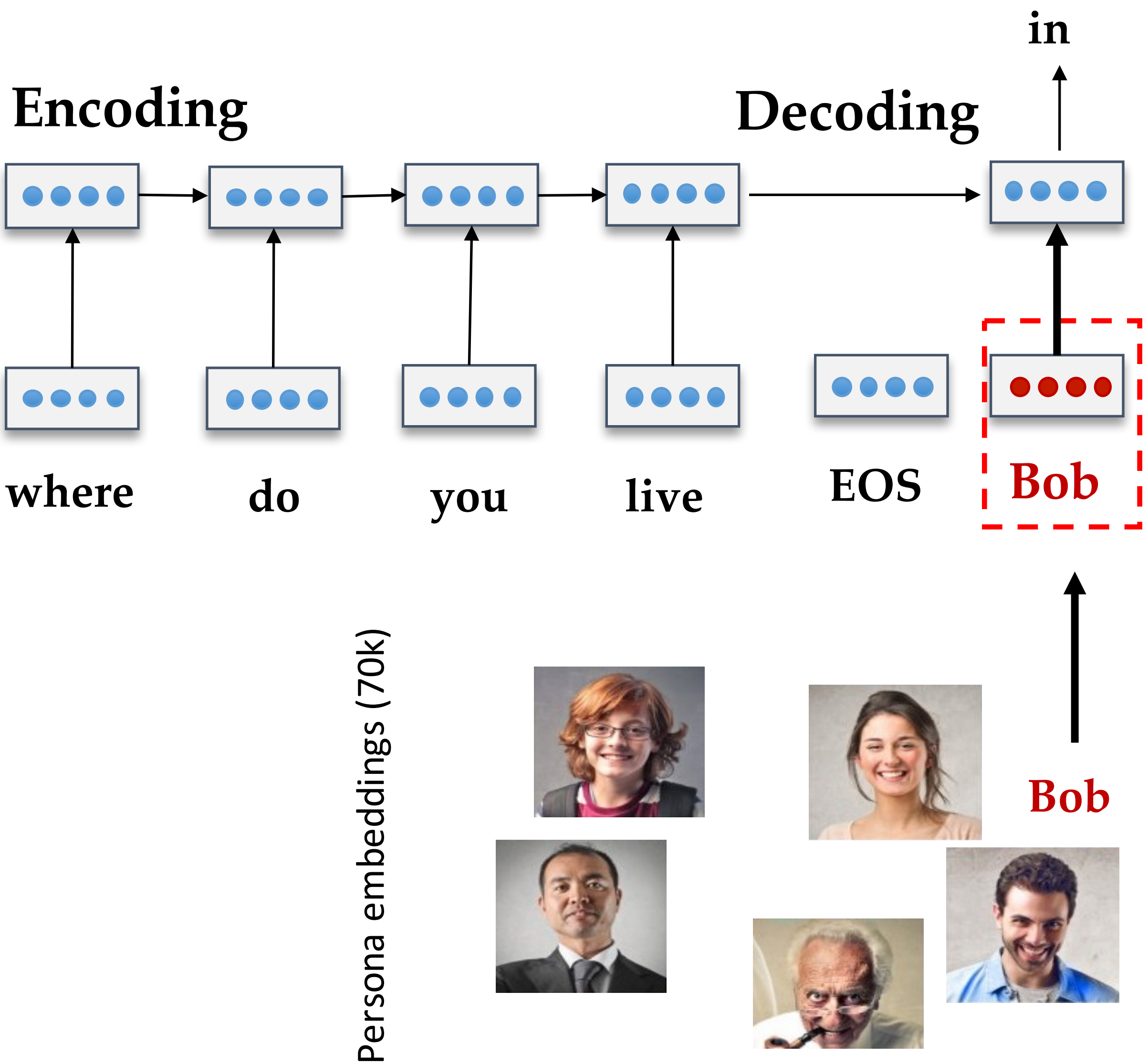

# Persona seq2seq model



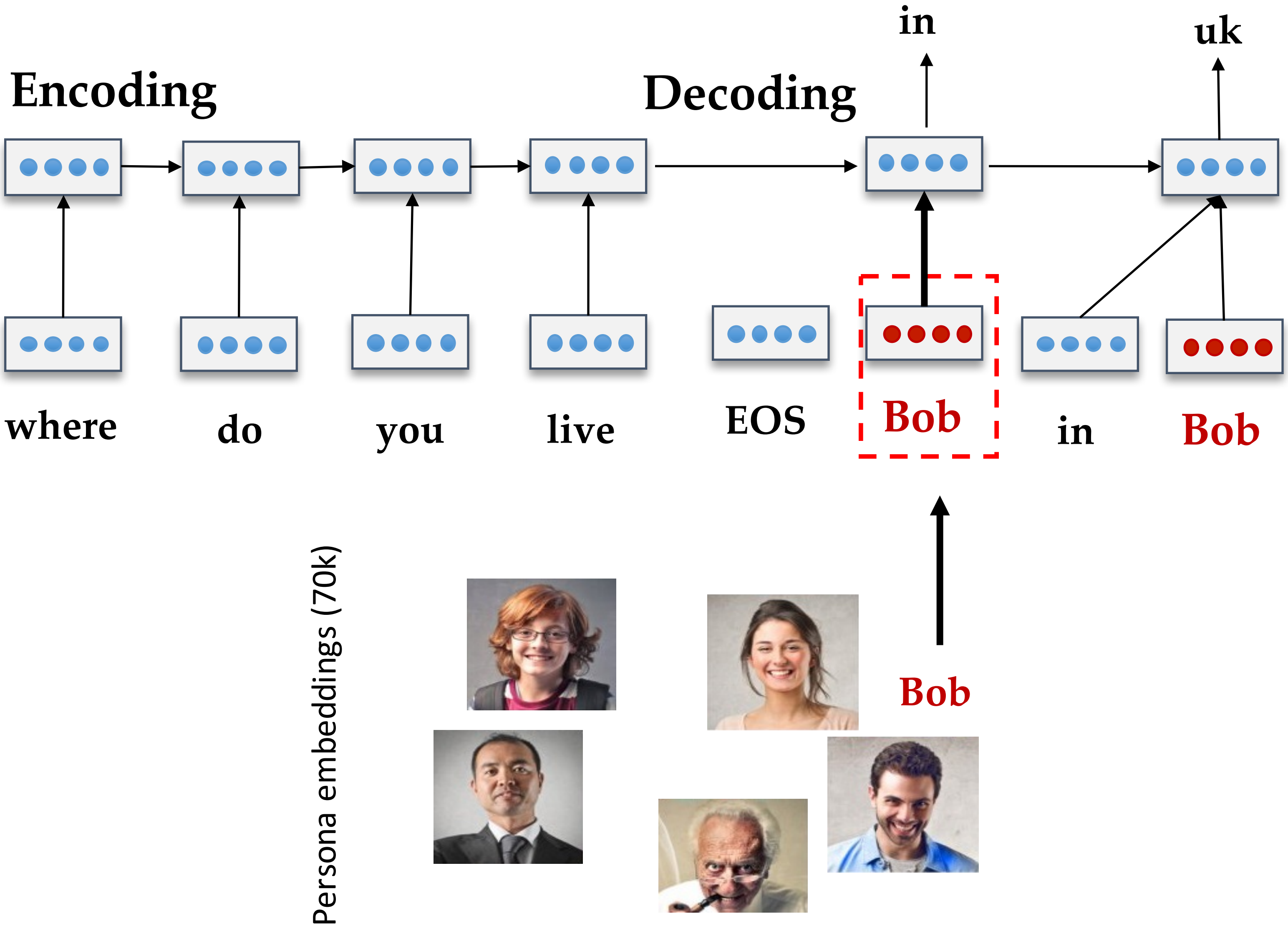
# Persona seq2seq model



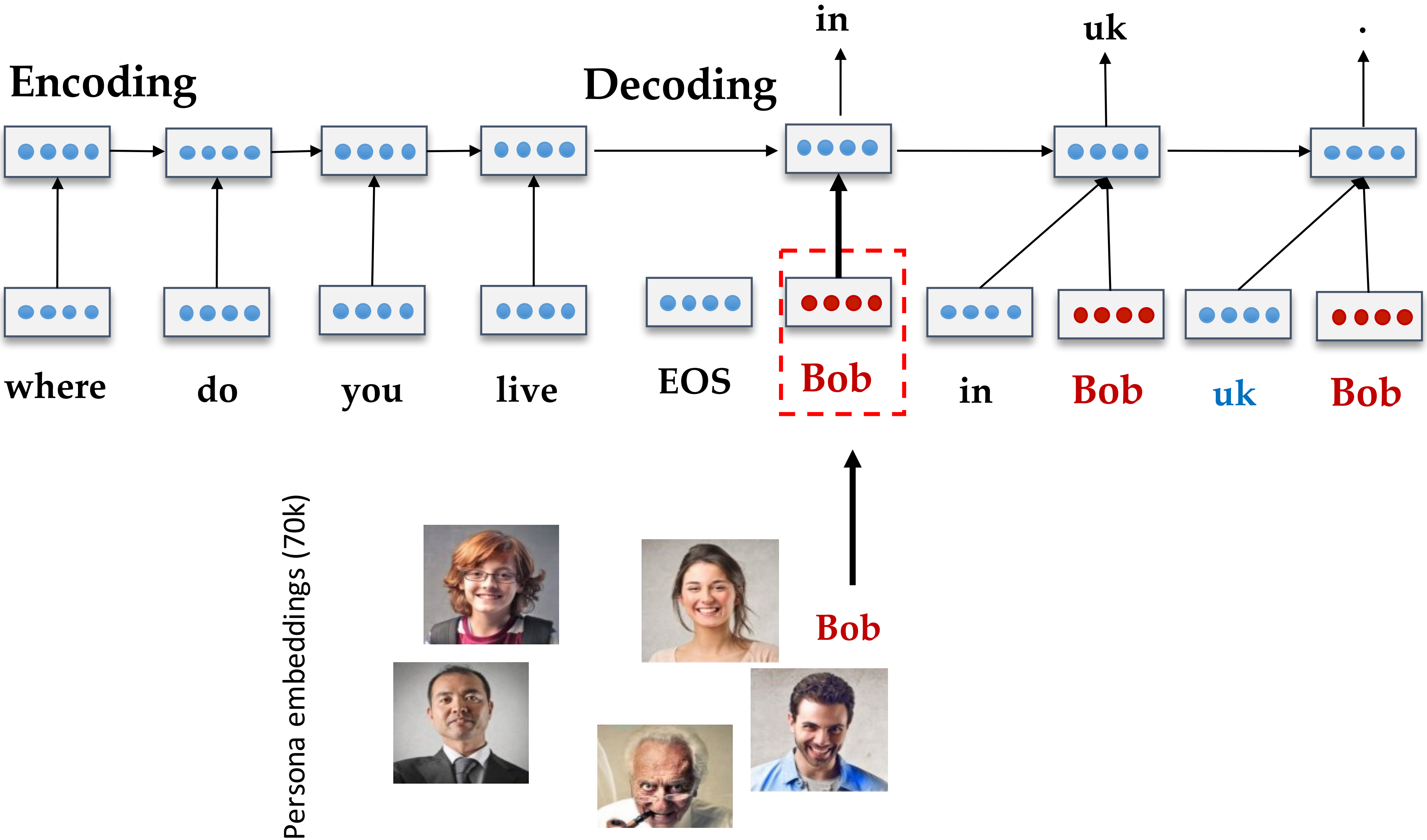
# Persona seq2seq model



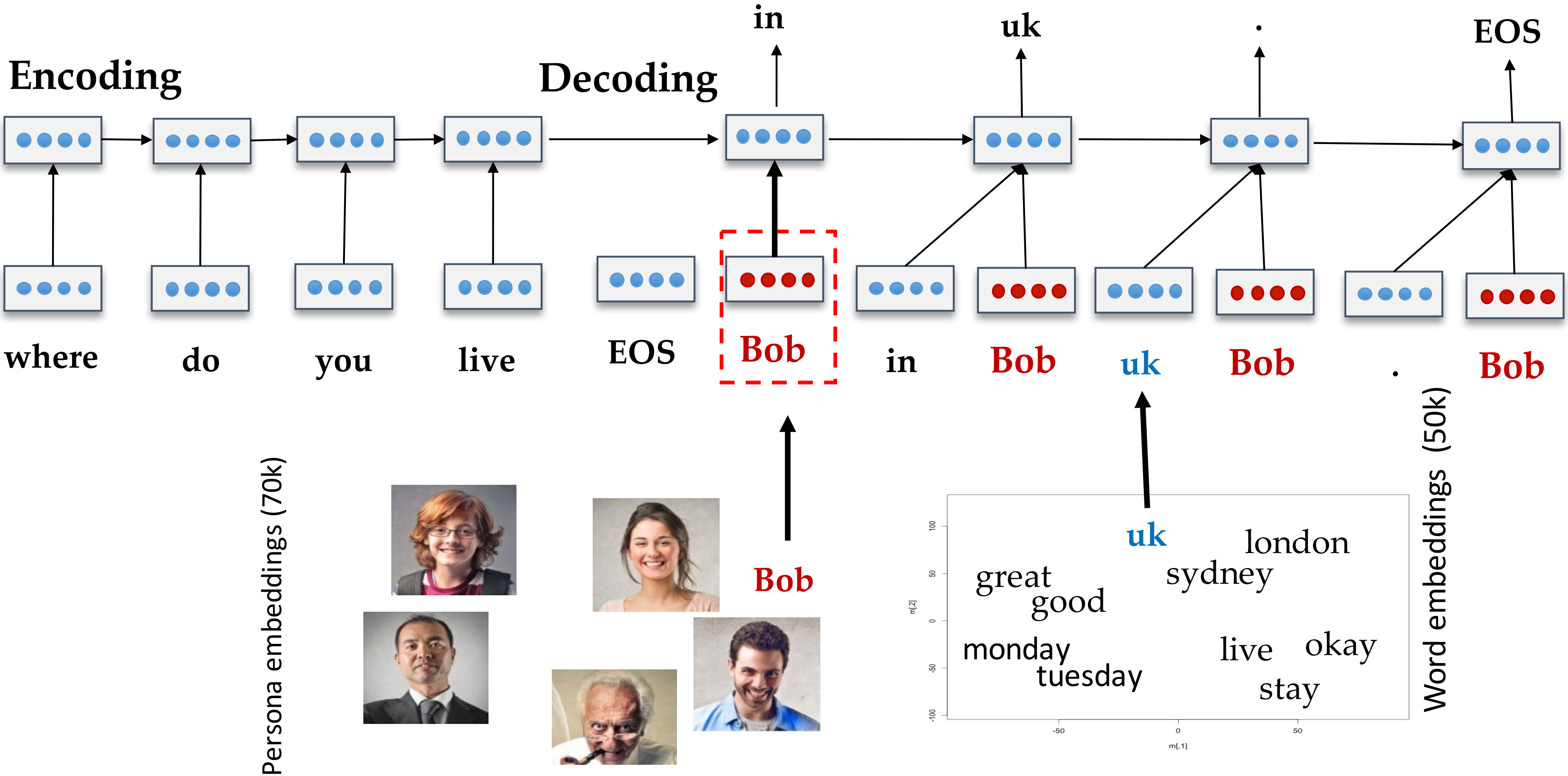
# Persona seq2seq model



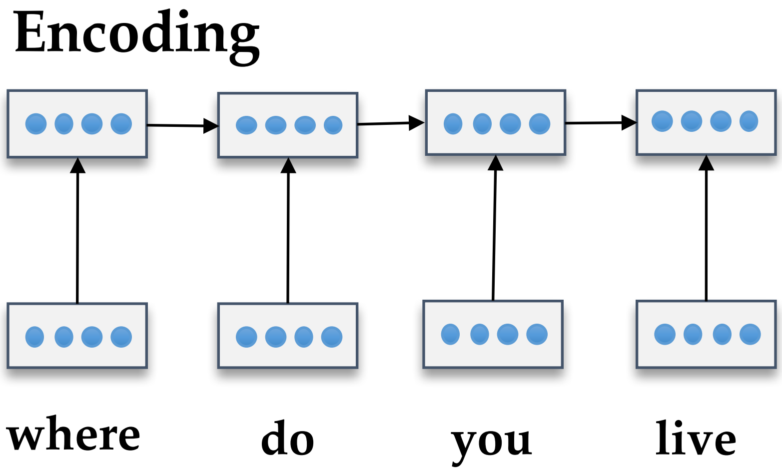
# Persona seq2seq model



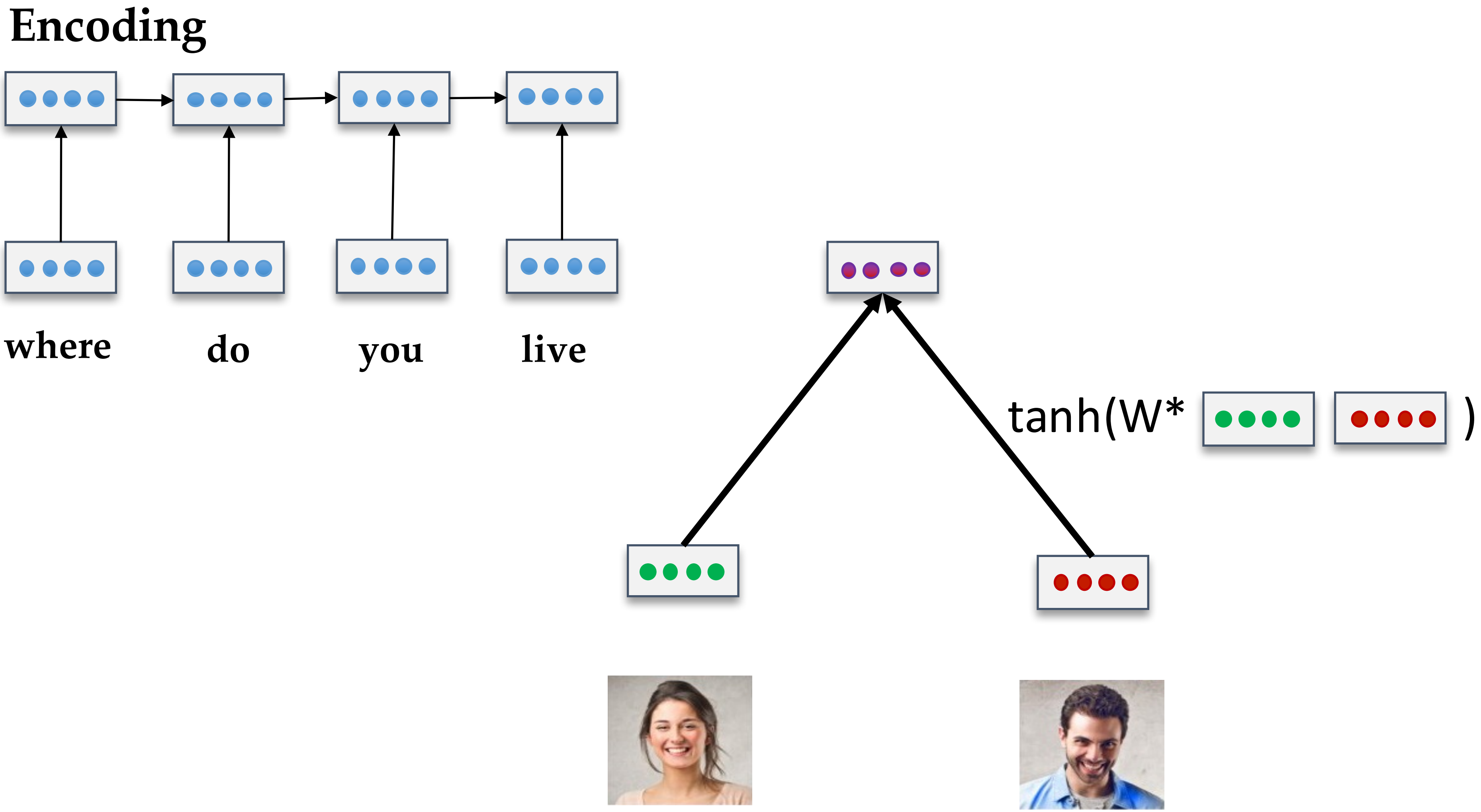
# Persona seq2seq model



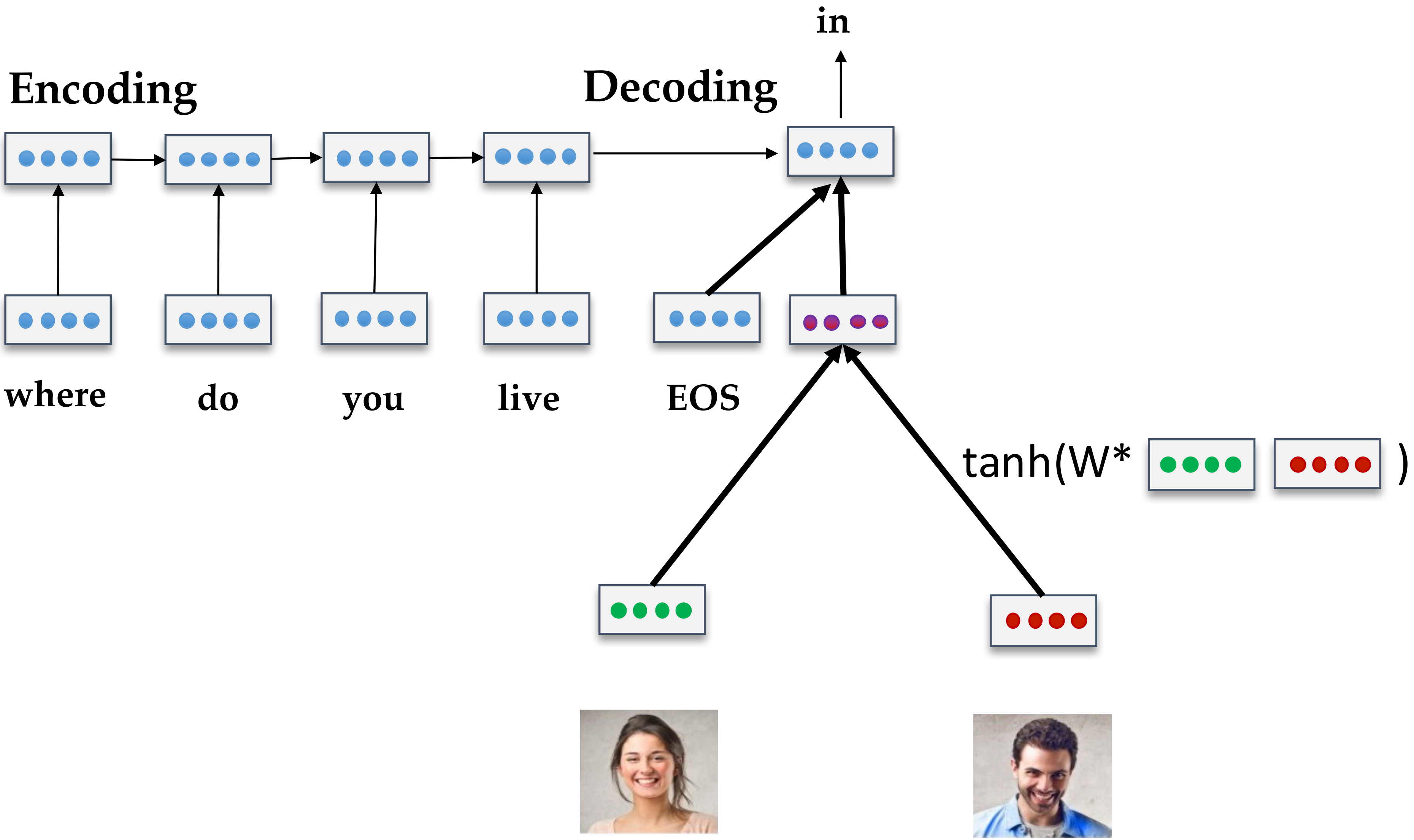
# Interaction Seq2Seq model



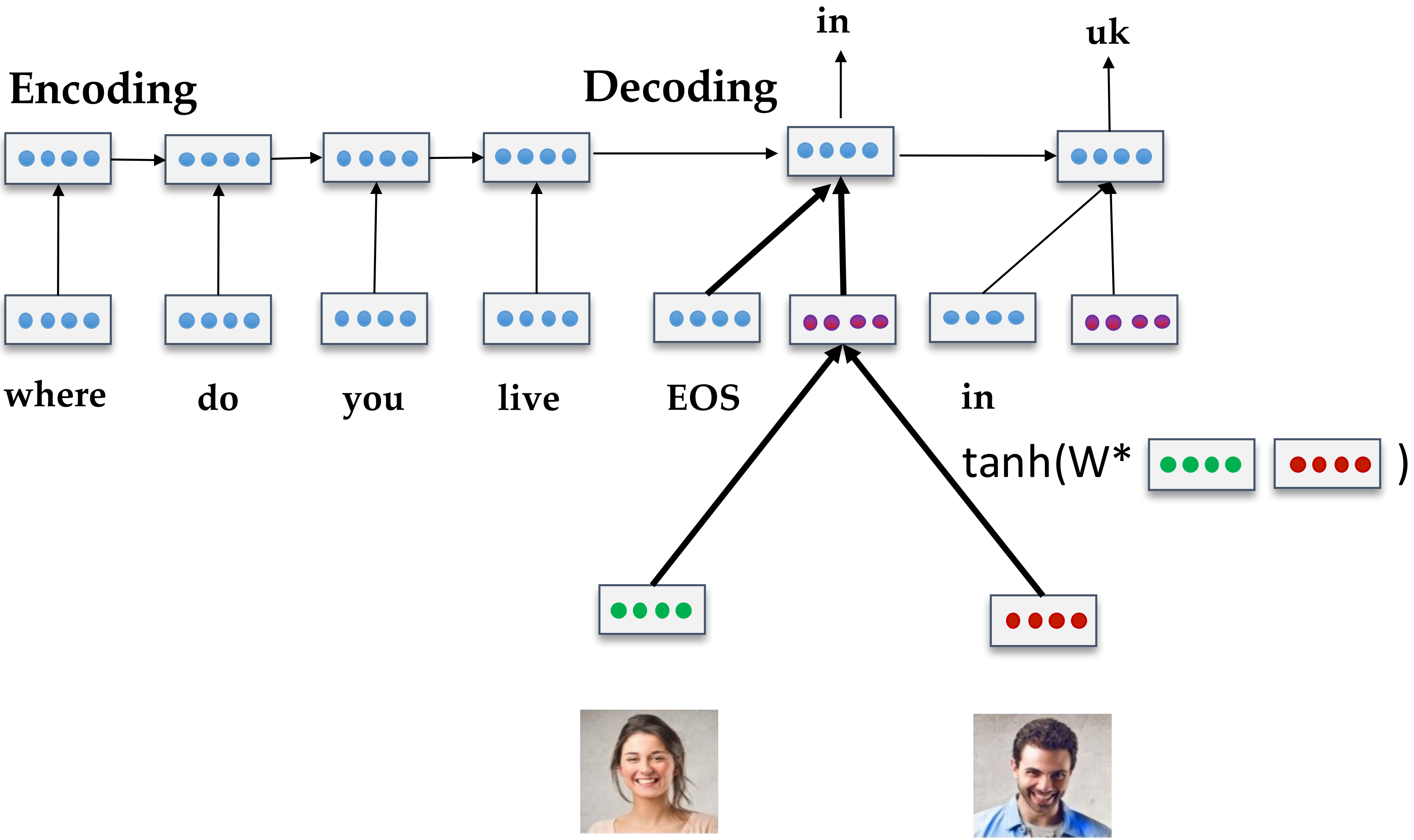

# Interaction Seq2Seq model



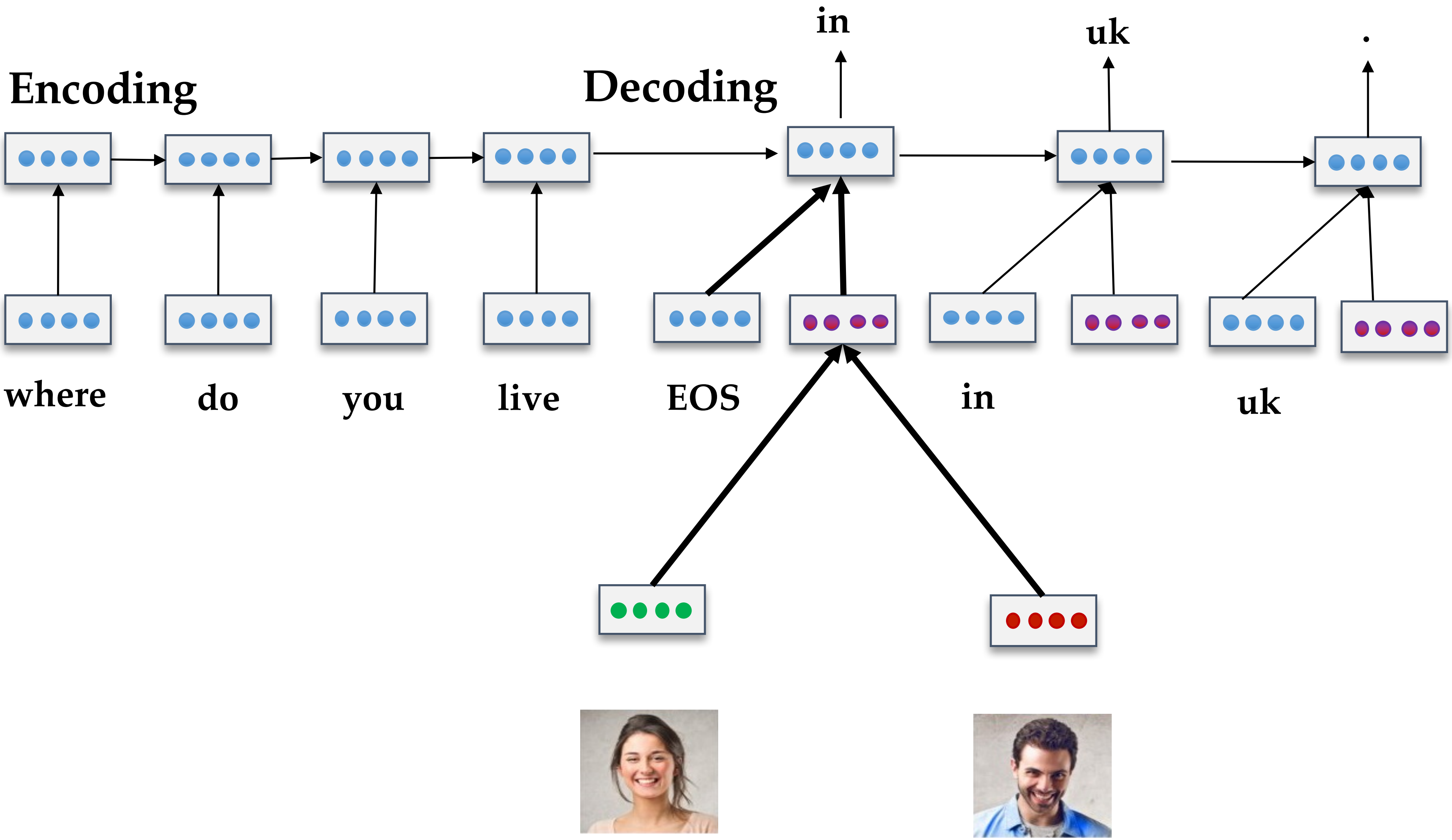
# Interaction Seq2Seq model



# Interaction Seq2Seq model



# Interaction Seq2Seq model



# 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

| | Seq2Seq | Speaker Model |
|---|---|---|
| Perplexity | 47.2 | 42.2 (-10.6%) |
| BLEU (without MMI) | 0.92 | 1.12 (+21.7%) |
| BLEU (with MMI) | 1.41 | 1.66 (+11.7%) |

# 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 beaf ?

# 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


| | Seq2Seq Model | Persona Model |
|---|---|---|
| Item1 | +1 | |
| Item2 | | +1 |

# Human Evaluation

| Seq2Seq Model | Persona Model |
|---|---|
| 0.84 | 1.33 (+34.7%) |

# 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? |
| *response* | I live in England. |
| *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. |

# 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? |
| *response* | I live in England. |
| *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. |

# 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? |
| *response* | I live in England. |
| *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. |

# 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? |
| *response* | I live in England. |
| *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. |

# 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? |
| *response* | I live in England. |
| *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. |

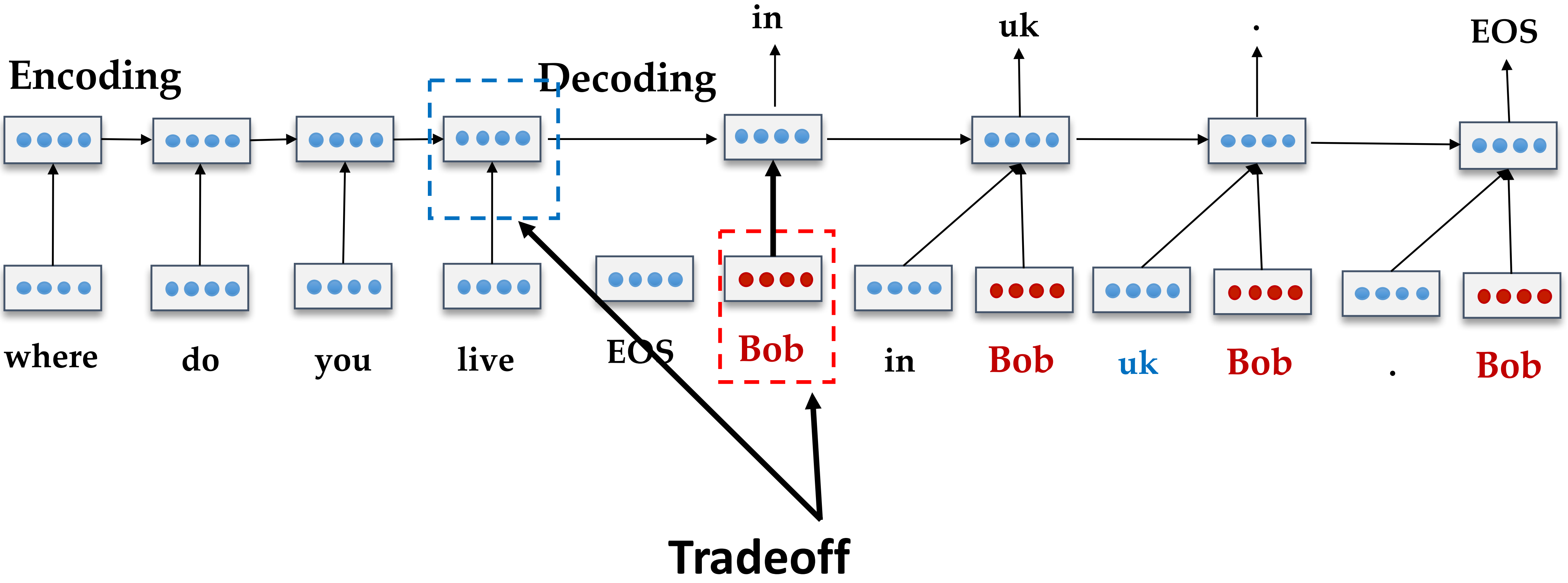
| User2 | |
|---|---|
| *message* | Where is your hometown? |
| *response* | I live in Indonesia. |
| *message* | Where are you from? |
| *response* | England, you? |
| *message* | In which city do you live now? |
| *response* | I live in Jakarta. |
| *message* | In which country do you live now? |
| *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. |

# 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? |
| *response* | I live in England. |
| *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. |

# Persona seq2seq model



# 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

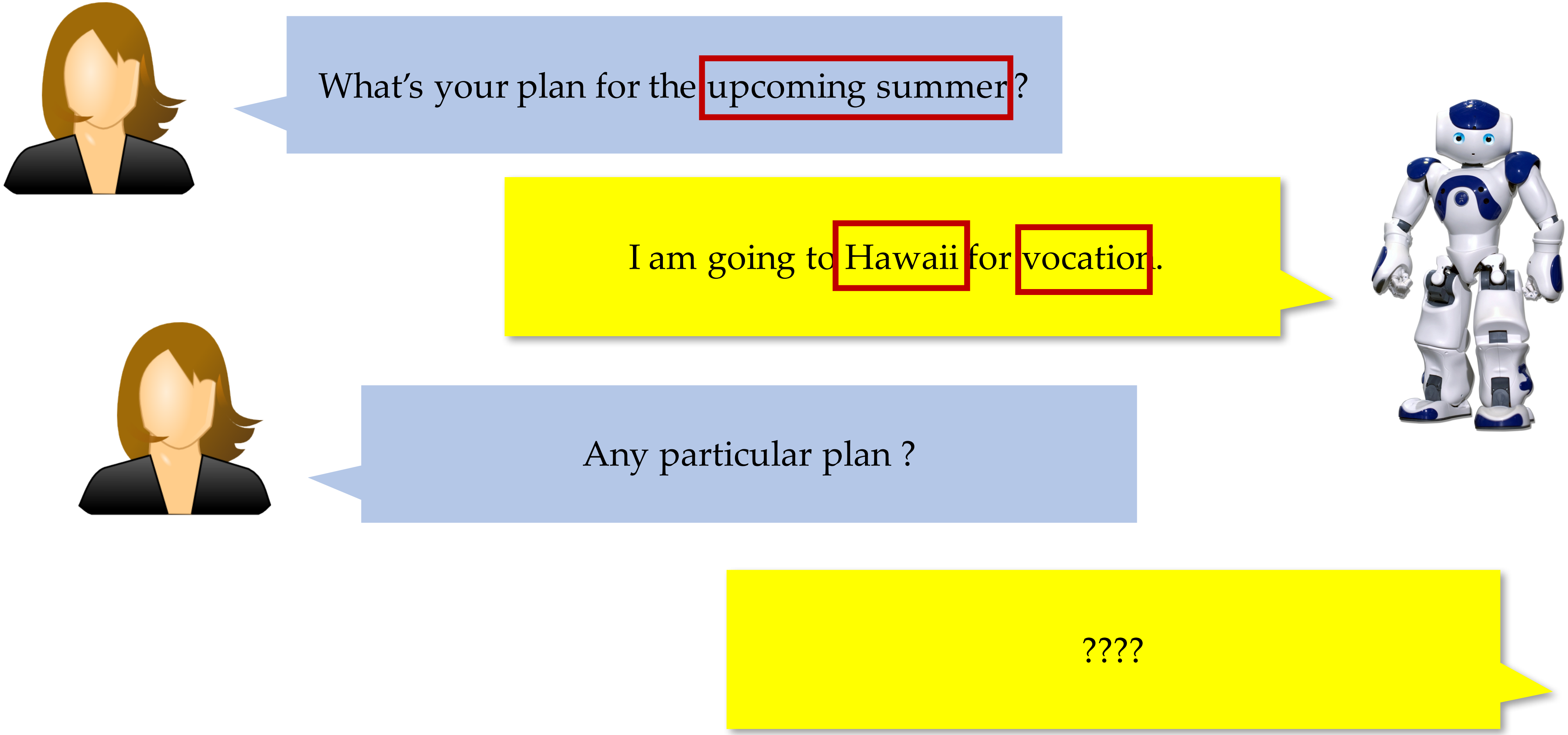
What’s your plan for the upcoming summer ?

I am going to Hawaii for vocation.

Any particular plan ?

????

# Multi-context Response Generation



# 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**

What's your plan for the upcoming summer ?

Context c1

Context c2

I am going to Hawaii for vocation.

…

Any particular plan ?

Message: m

Response r

????

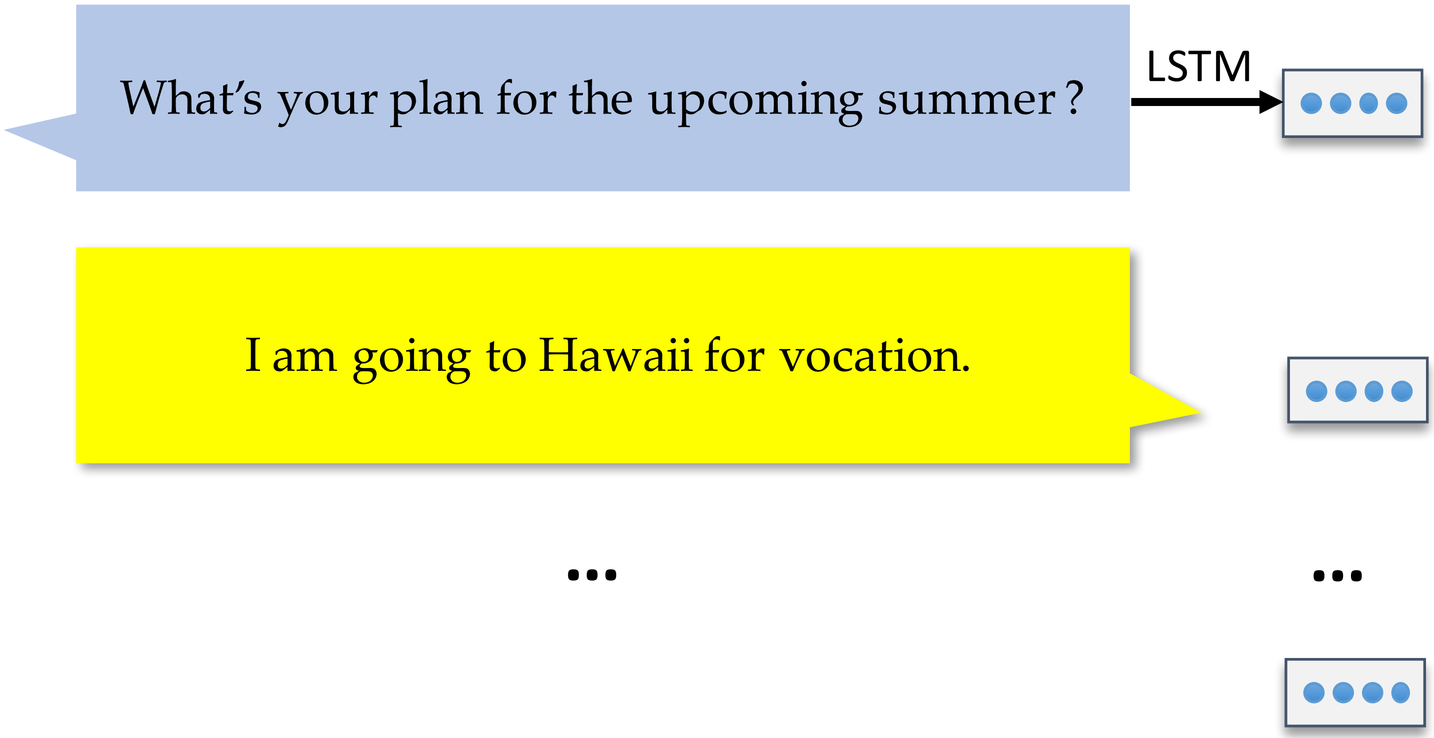
# Multi-context Response Generation

What's your plan for the upcoming summer ?

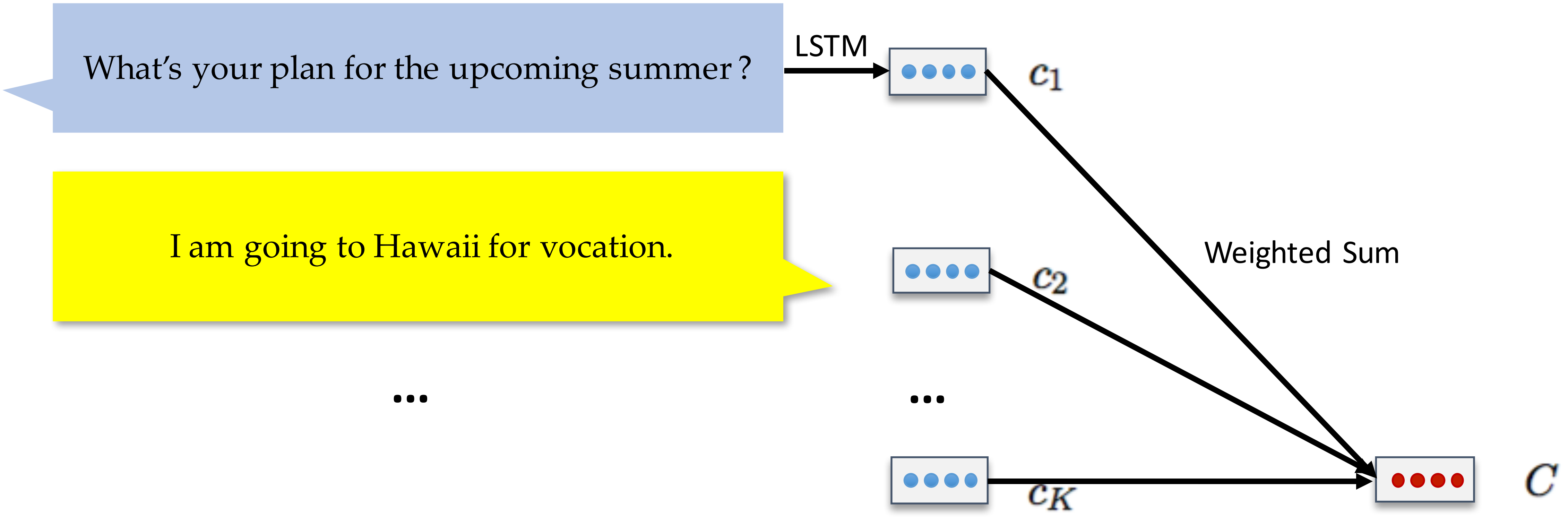
I am going to Hawaii for vocation.

…

# Multi-context Response Generation

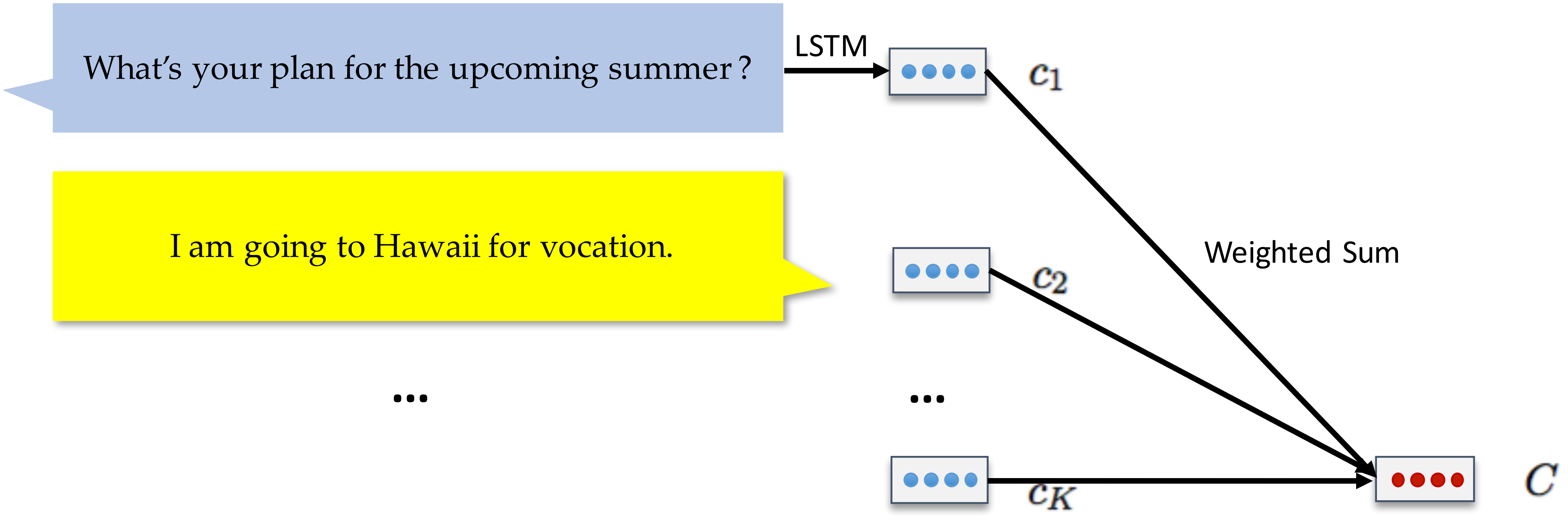


# Multi-context Response Generation



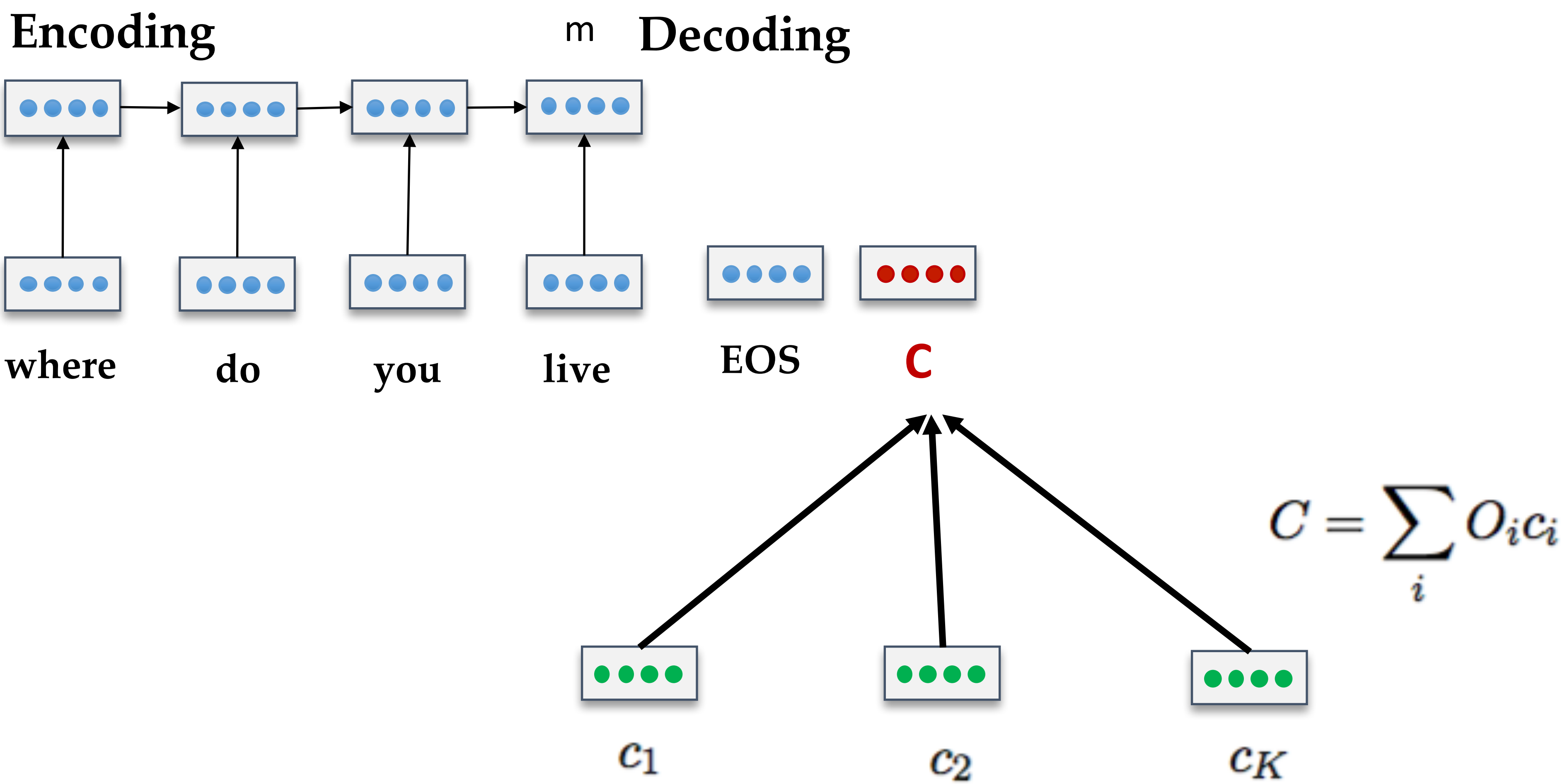

$$C = \sum_i O_i c_i$$

# Multi-context Response Generation

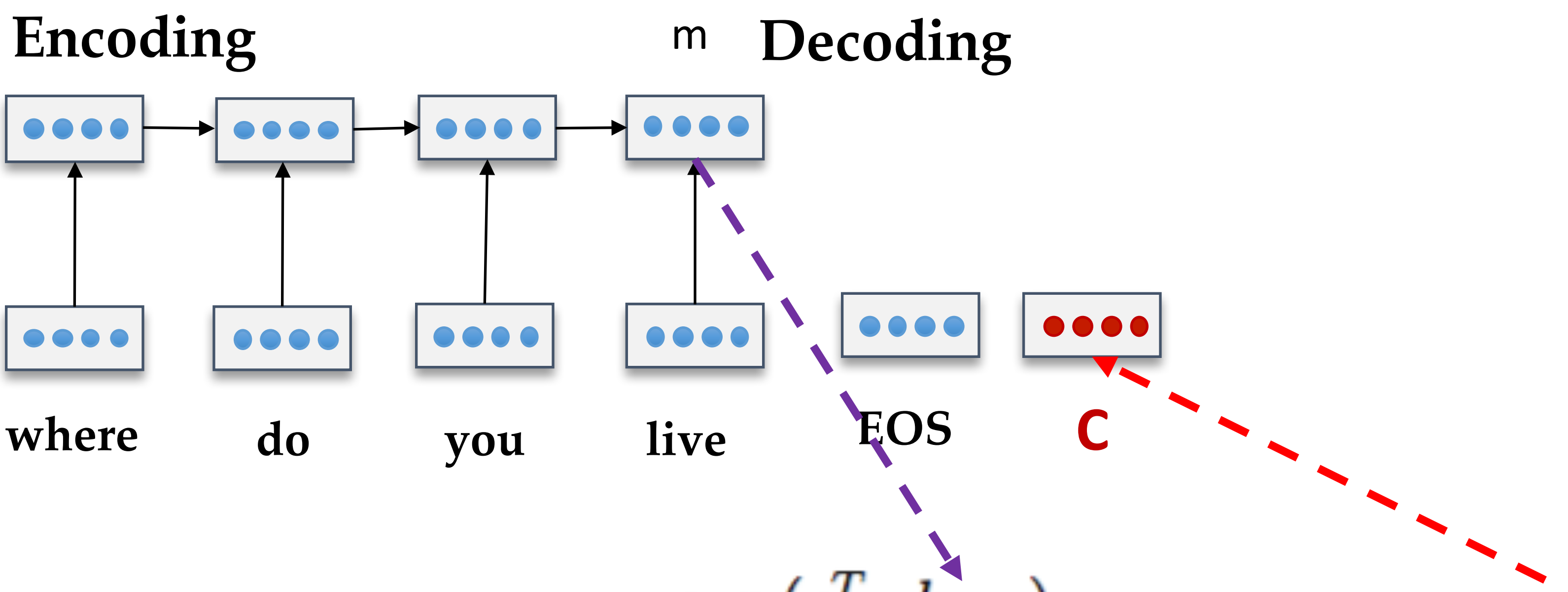



$$C = \sum_i O_i c_i$$ Memory Network (Weston et al., 2014)

# Multi-context Response Generation



# Multi-context Response Generation

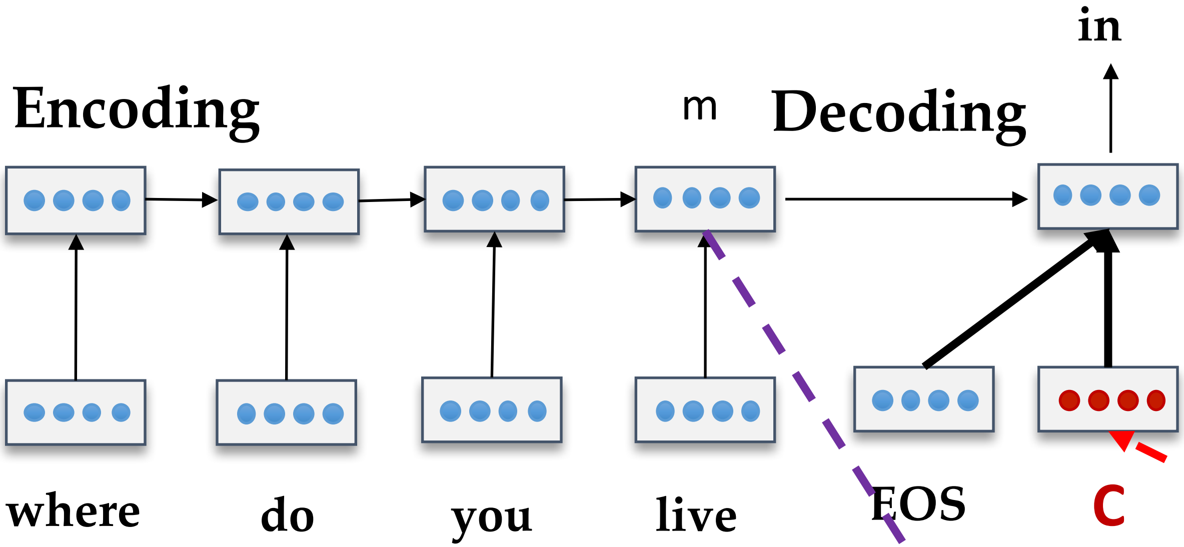



$$O_i = \frac{\exp\left(c_i^T \cdot h_{t-1}\right)}{\sum_i \exp\left(c_i^T \cdot h_{t-1}\right)}$$

$$C = \sum_i O_i c_i$$

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

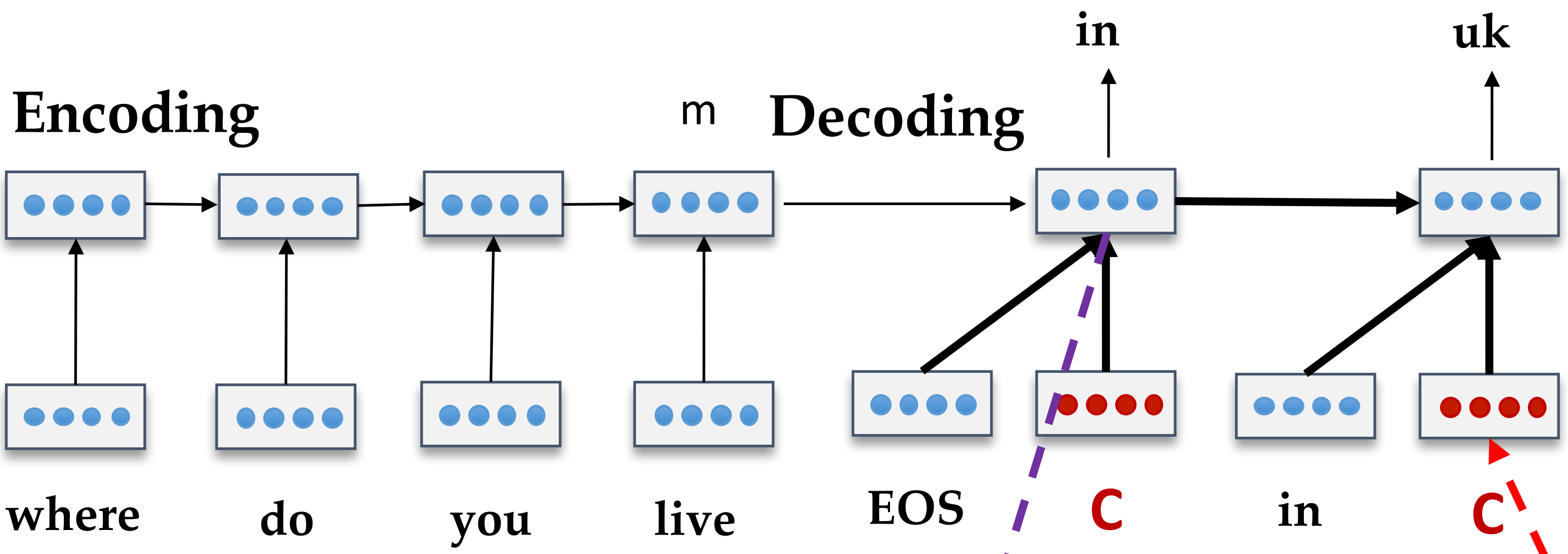
# Multi-context Response Generation




$$O_i = \frac{\exp\left(c_i^T \cdot h_{t-1}\right)}{\sum_i \exp\left(c_i^T \cdot h_{t-1}\right)}$$

$$C = \sum_i O_i c_i$$

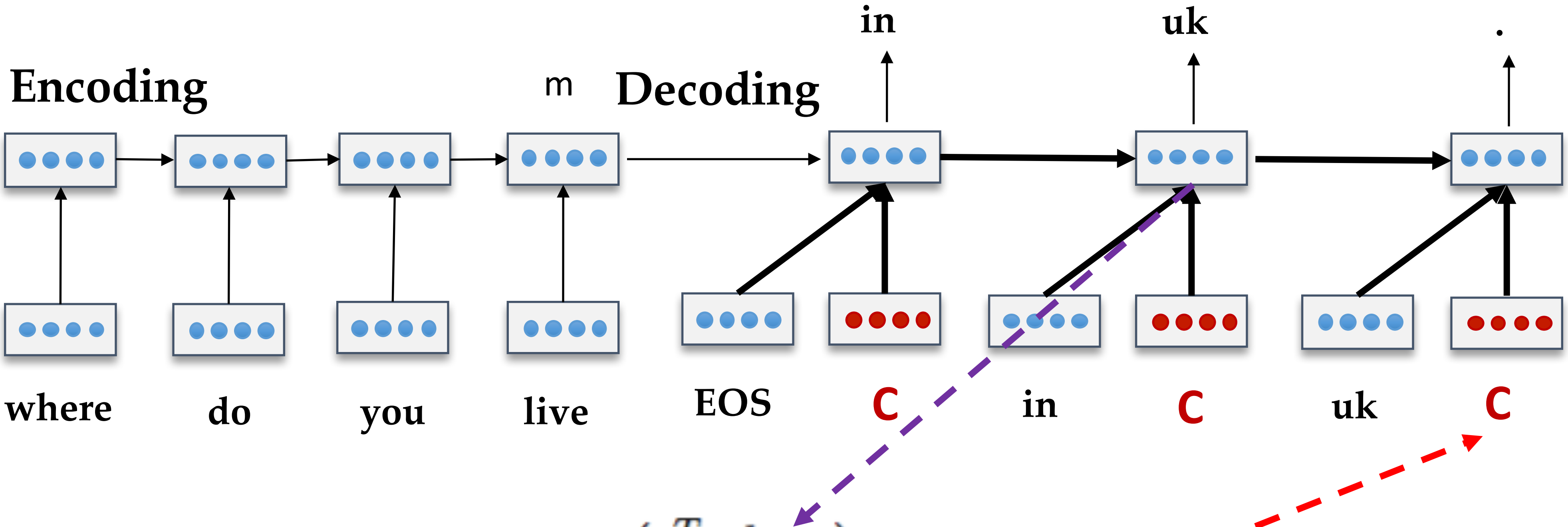
# Multi-context Response Generation




$$O_i = \frac{\exp\left(c_i^T \cdot h_{t-1}\right)}{\sum_i \exp\left(c_i^T \cdot h_{t-1}\right)}$$

$$C = \sum_i O_i c_i$$

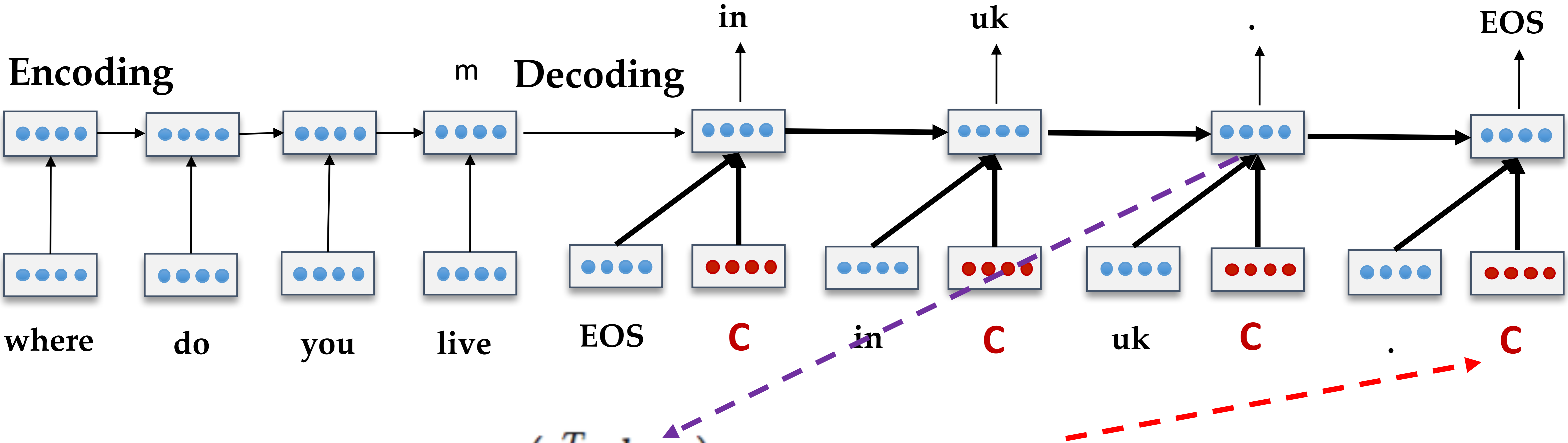
# Multi-context Response Generation




$$O_i = \frac{\exp\left(c_i^T \cdot h_{t-1}\right)}{\sum_i \exp\left(c_i^T \cdot h_{t-1}\right)}$$

$$C = \sum_i O_i c_i$$

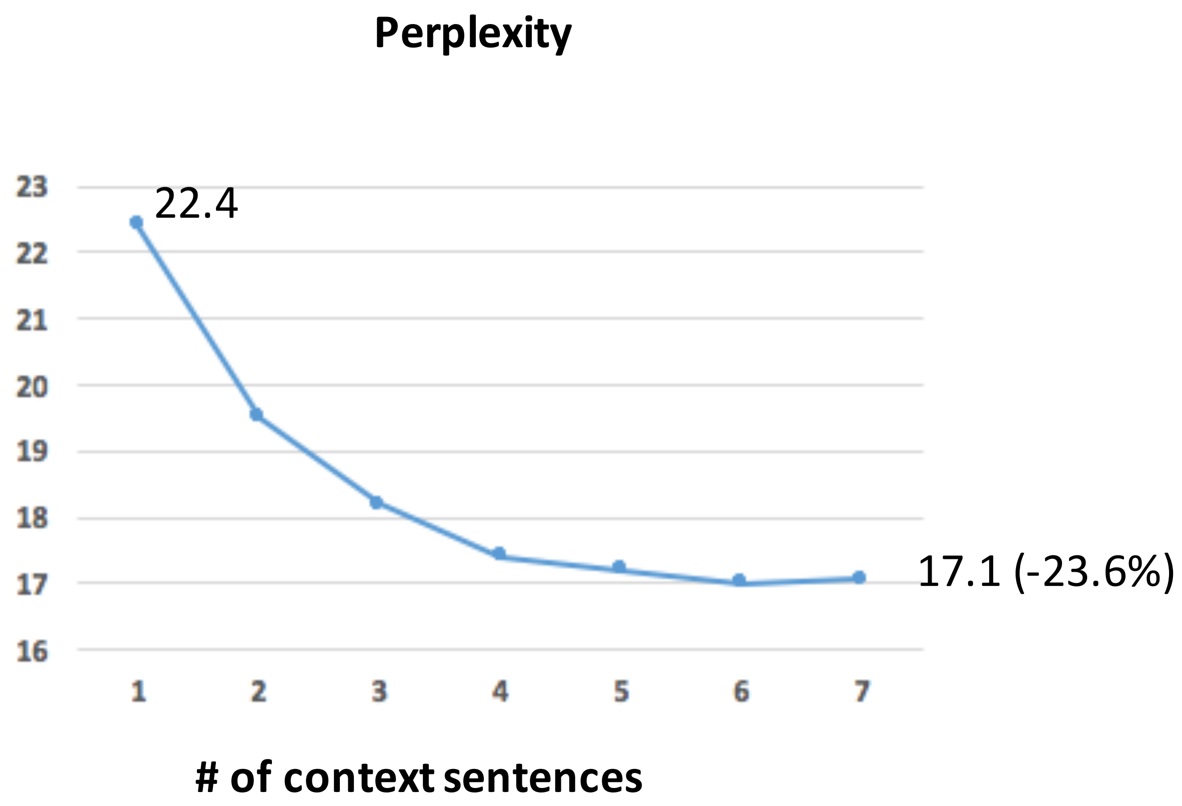
# Multi-context Response Generation



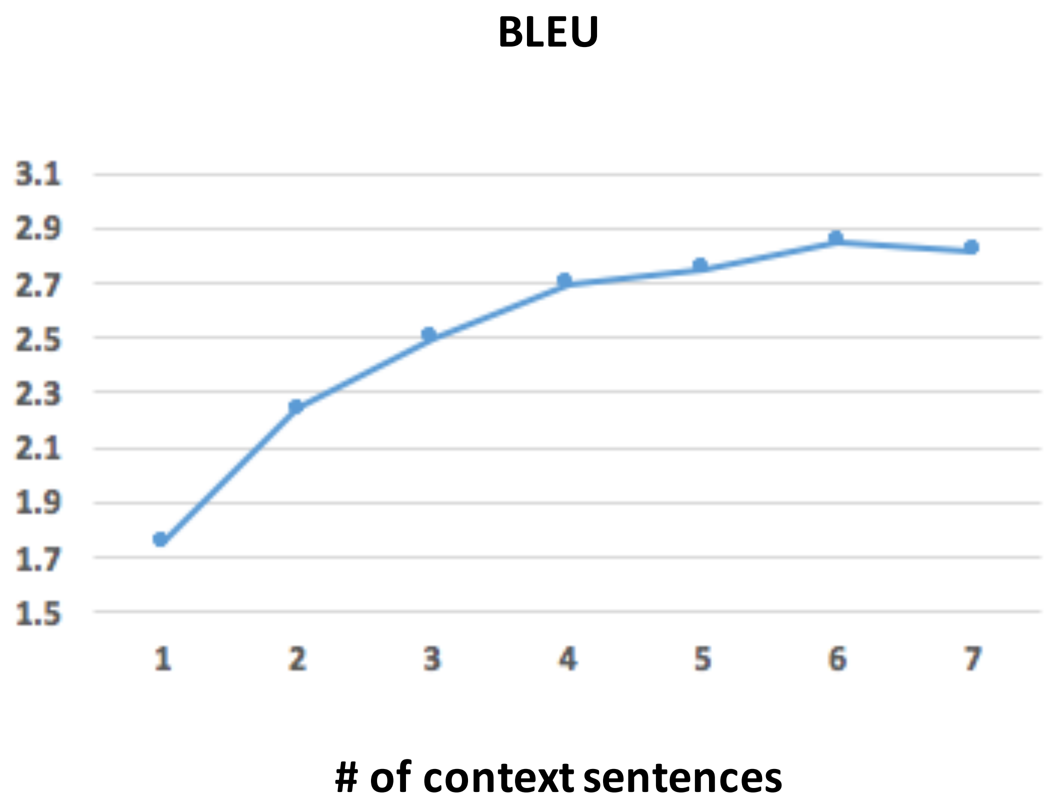

$$O_i = \frac{\exp\left(c_i^T \cdot h_{t-1}\right)}{\sum_i \exp\left(c_i^T \cdot h_{t-1}\right)}$$

$$C = \sum_i O_i c_i$$

# Results on the Opensubtitle Dataset



# Results on the Opensubtitle Dataset



# 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

$$\neq$$

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

# Supervised Learning

# Supervised Learning

Data + Labels/Rewards

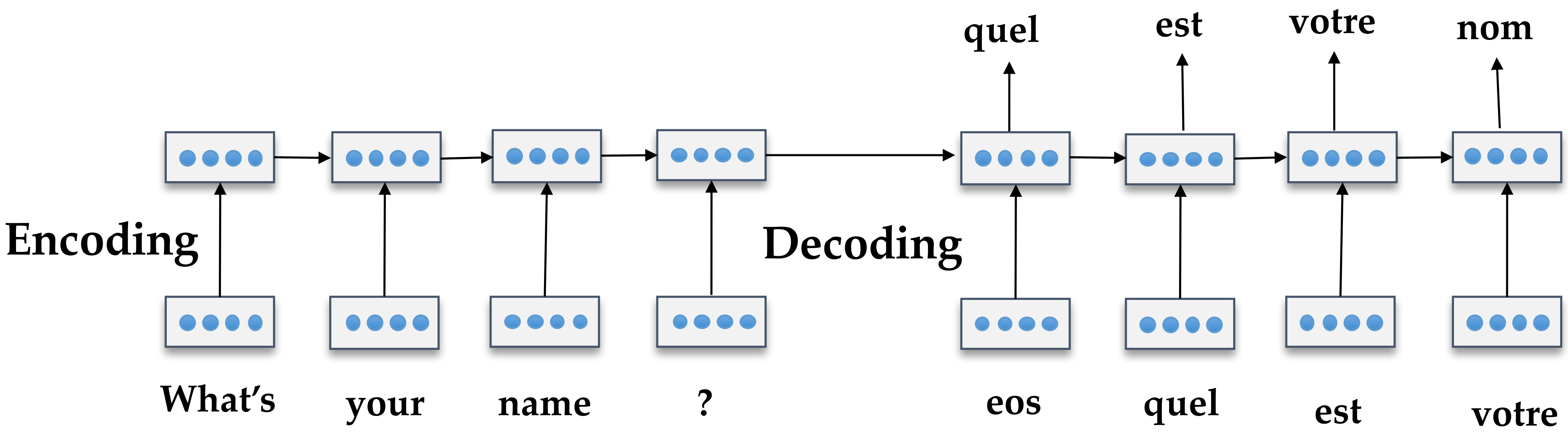
# Supervised Learning

| Data | Labels |
|---|---|
| You are a good boy. |  |
| You are a bad boy. |  |

# Supervised Learning

Data

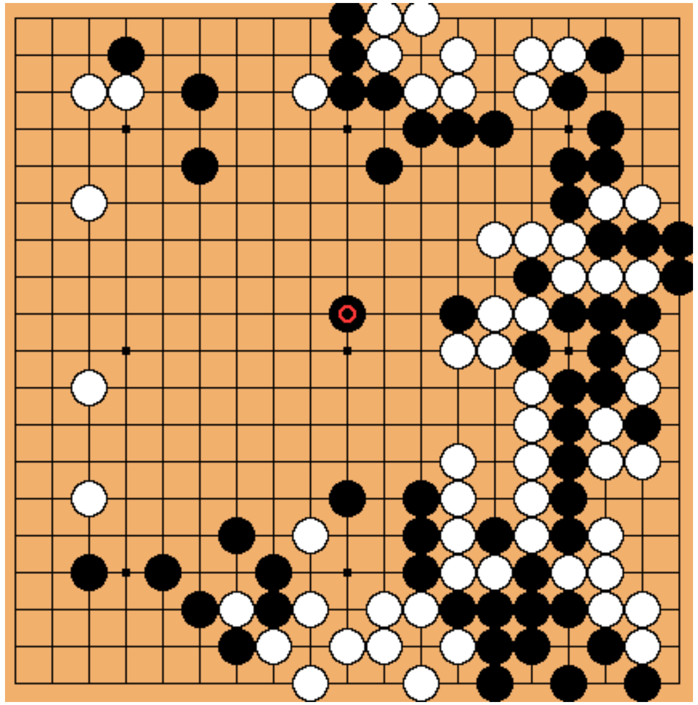
Labels



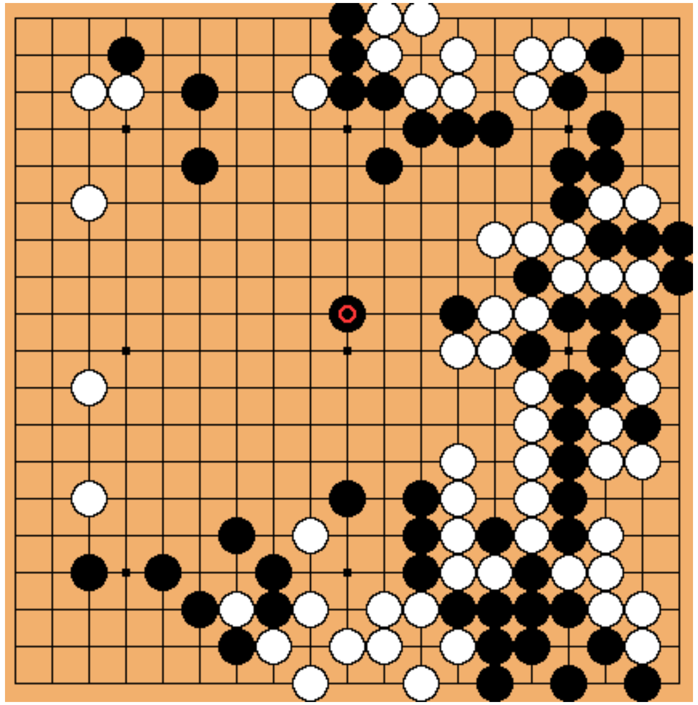
# What if labels or rewards is not immediate clear ?

# 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 ?

i 'm 16 .

16 ?

# Simulation

How old are you ?

i 'm 16 .

16 ?

i don 't know what you 're talking about

# 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

# 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

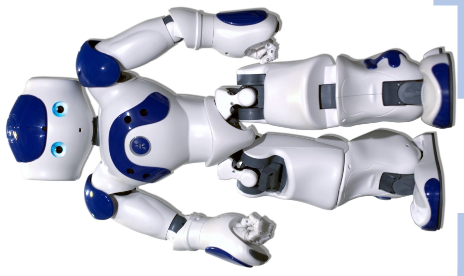
# Simulation

How old are you ?

i 'm 16 .

16 ?

i don 't know what you 're talking about



you don 't know what you 're saying

i don 't know what you 're talking about

you don 't know what you 're saying

# 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

you don 't know what you 're saying

# Simulation

**Bad Action**

How old are you ?

i 'm 16 .

16 ?

i don 't know what you 're talking about

you don 't know what you 're saying

i don 't know what you 're talking about

you don 't know what you 're saying

# 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 ...

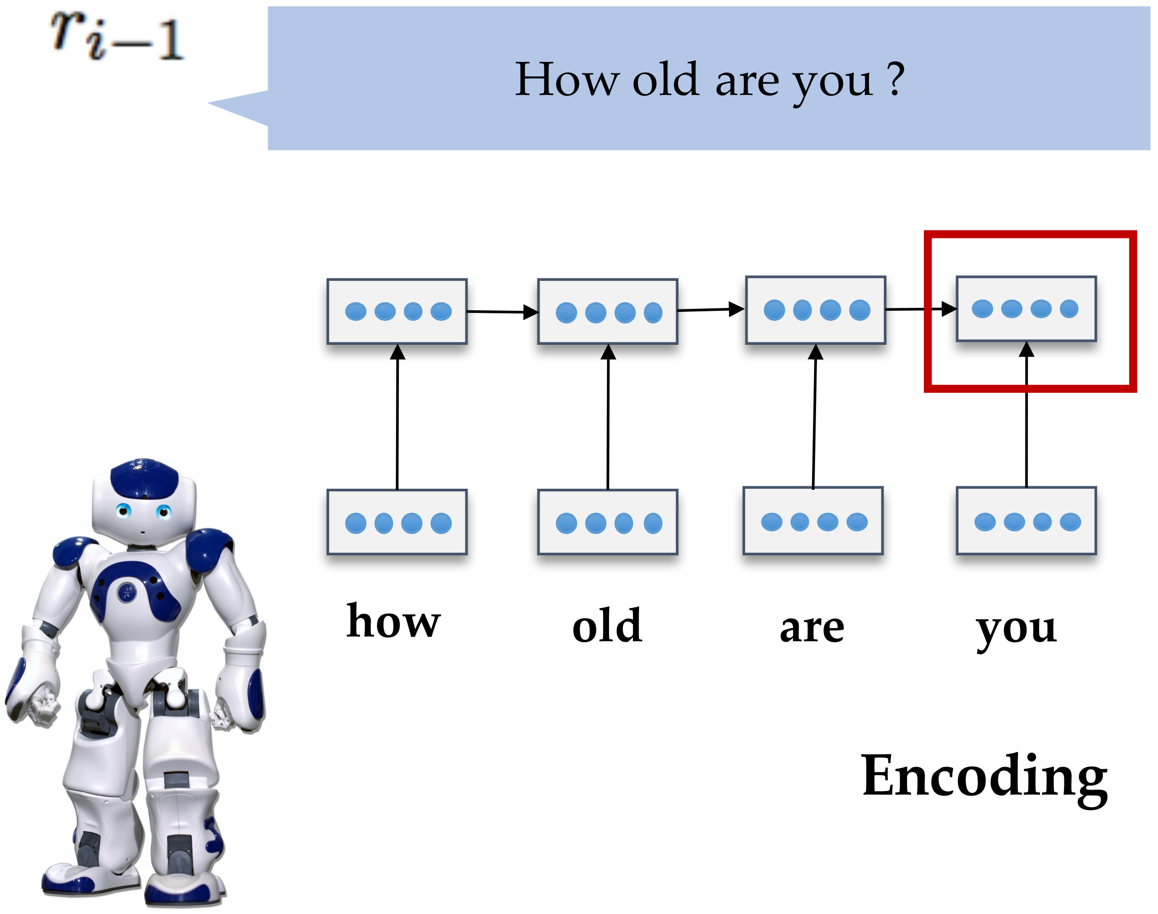
# Can RL (to some extent) achieve this goal?

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

$r_{i-1}$



# Notations: Action

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

tf-idf

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

tf-idf

Informativeness

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

tf-idf

# of turns it takes before generating dull responses

# Notations: Reward

$r_{i-1}$

How old are you ?

$r_i$

i 'm 16 .

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

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

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

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 $r_i$ Infinite #

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 $\mathbb{A}$

Action: Generating current response $r_i$

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

Approximation1: Sample a small list of candidates $\mathbb{A}$

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

Normalization Part

**Future Reward** $\hat{R}(r_i)$

$$\hat{R}(r_i) = \underbrace{\boxed{R(r_i)}}_{\text{Immediate Reward}} + \gamma \sum_{r_{i+1}} p(r_{i+1}|r_i)\hat{R}(r_{i+1})$$

## Future Reward $\hat{R}(r_i)$

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

Immediate Reward

$$\mathrm{MMI}(r_i, r_{i-1}) \cdot \mathrm{reward}(r_i)$$

**Future Reward** $\hat{R}(r_i)$

$$\hat{R}(r_i) = R(r_i) + \underbrace{\boxed{\gamma \sum_{r_{i+1}} p(r_{i+1}|r_i)\hat{R}(r_{i+1})}}_{\text{Future Reward}}$$

**Future Reward** $\hat{R}(r_i)$

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

Future Reward

$$\approx R(r_i) + \sum_{r_i \in \mathbb{A}} \frac{p(r_{i+1}|r_i)}{\sum_{t \in \mathbb{A}} p(r_{i+1}|r_{i1})} [\hat{R}(r_{i+1})]$$

# Simulation

# Simulation

# Simulation

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

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

A message from training set

# Simulation

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

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

A message from training set

**Decoding**

$r_{i,a}$

$r_{i,b}$

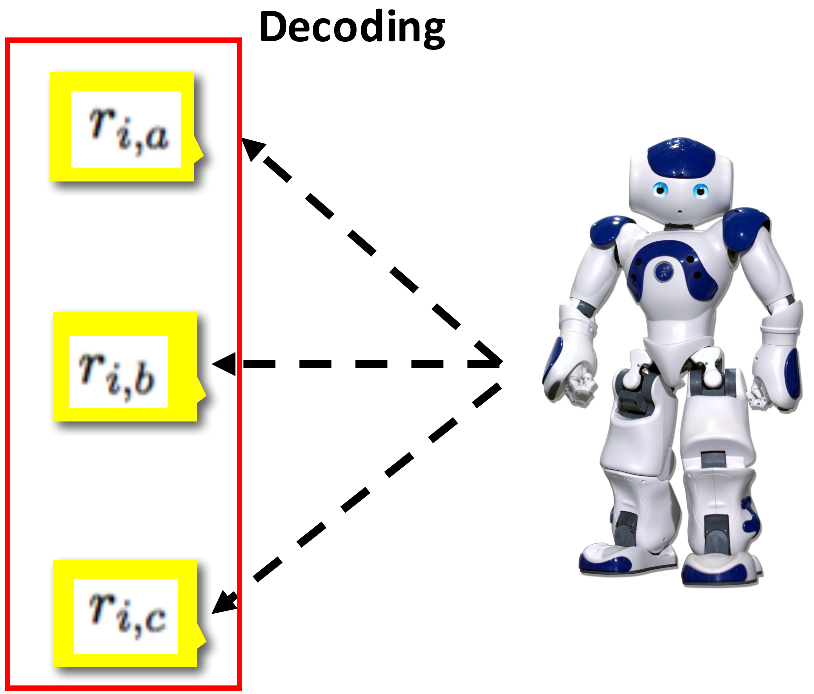
$r_{i,c}$

# Simulation

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

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

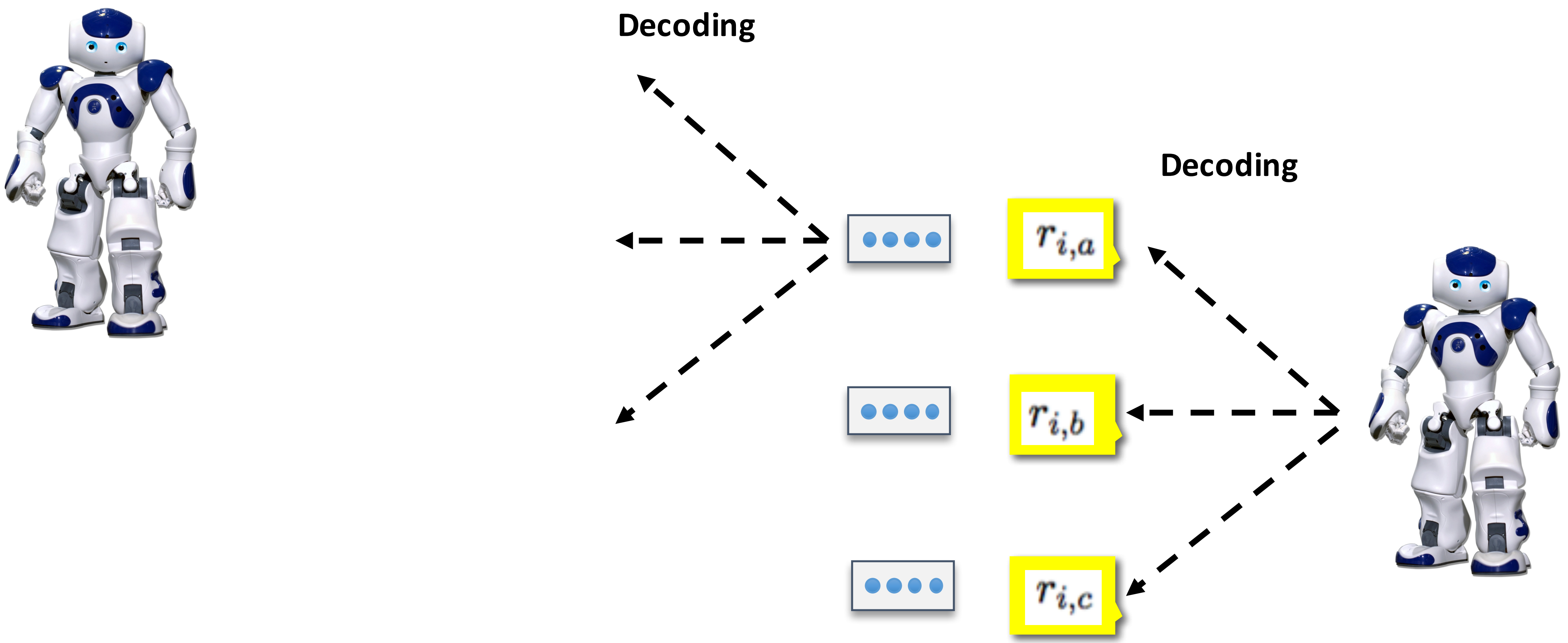
A message from training set



# Simulation

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

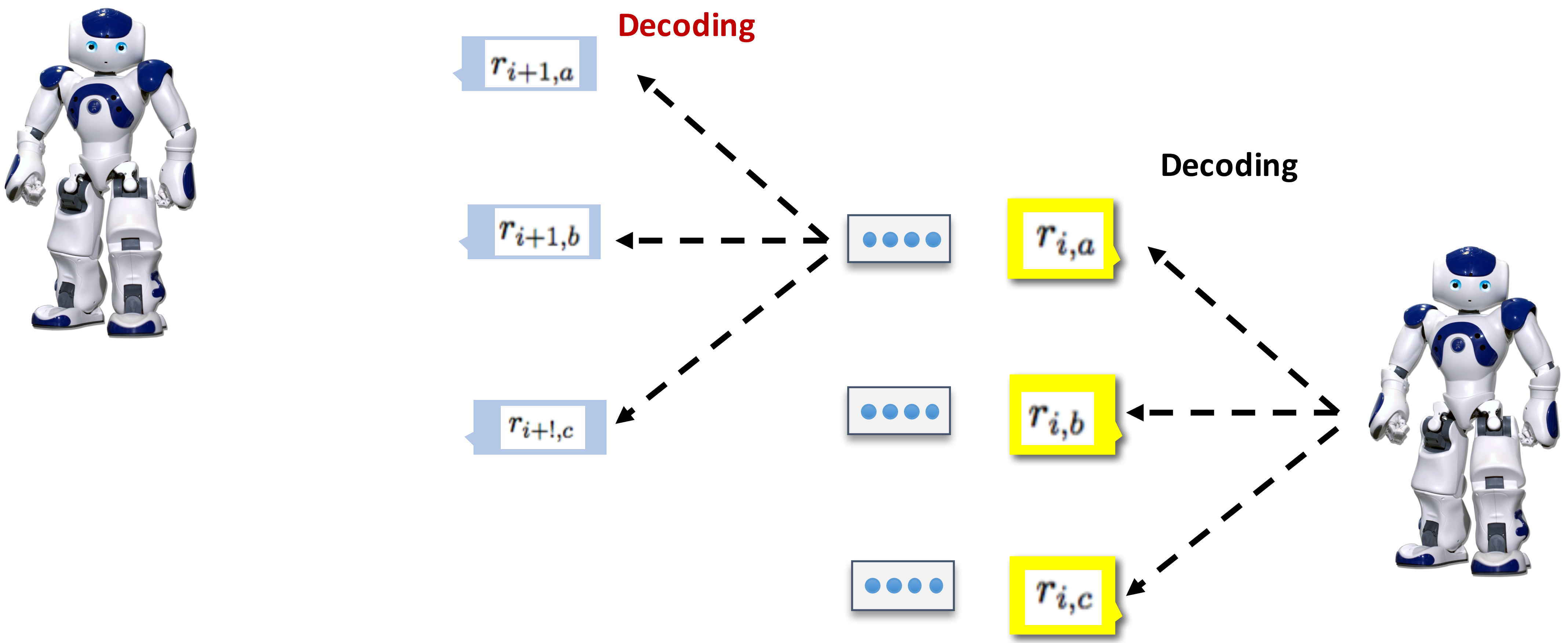
$$R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in \mathbb{A}} \frac{p(r_{i+1}|r_i)}{\sum_{t \in \mathbb{A}} p(r_{i+1}|r_i)}[R(r_{i+1})]$$



# Simulation

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

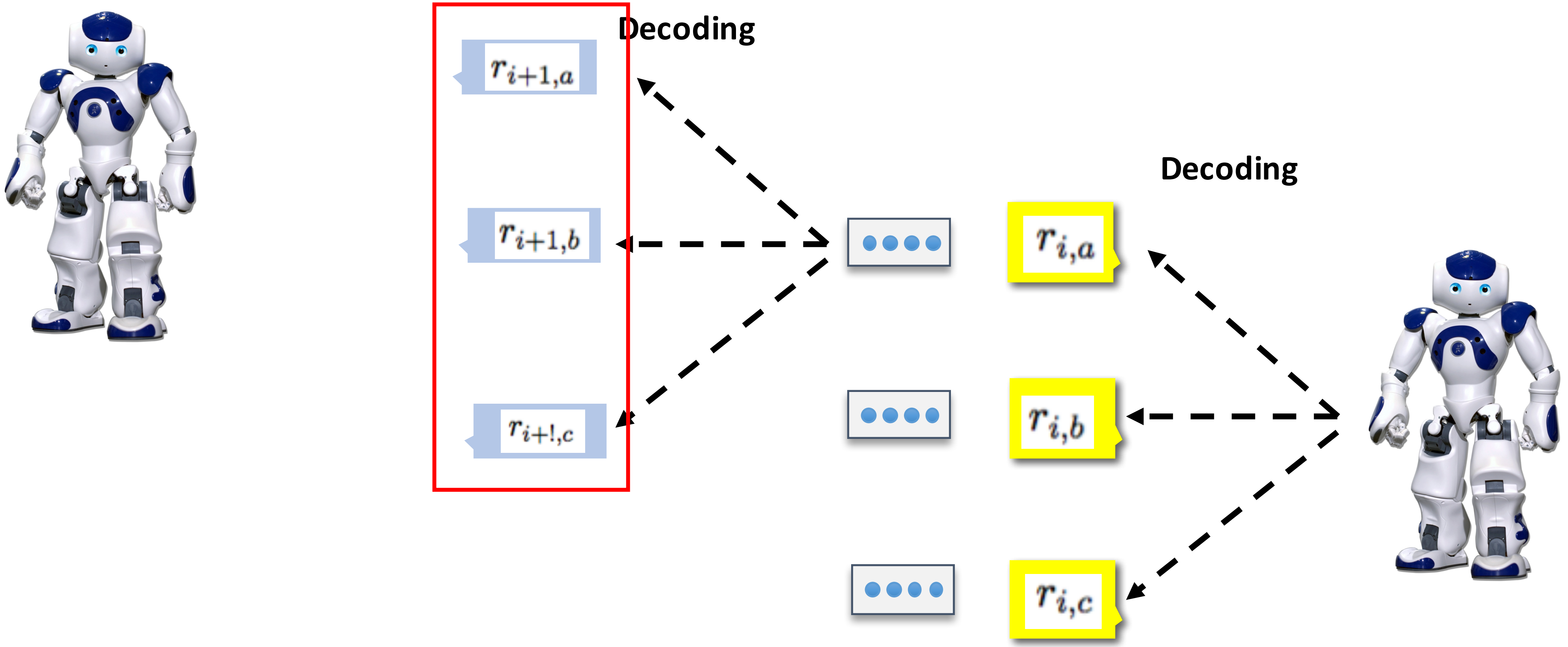
$$R(\hat{r}_i) \approx R(r_i) + \gamma \sum_{r_{i+1} \in \mathbb{A}} \boxed{\frac{p(r_{i+1}|r_i)}{\sum_{t \in \mathbb{A}} p(r_{i+1}|r_i)}} [R(r_{i+1})]$$

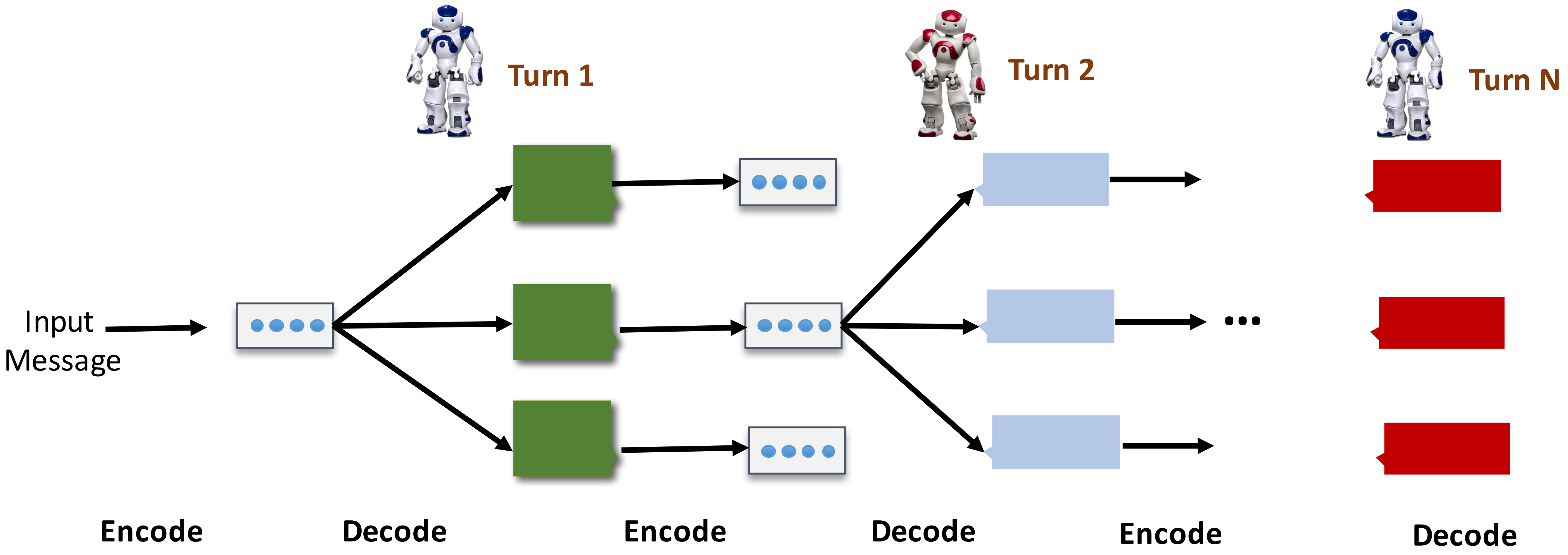


# Simulation

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

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




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

# Results

| Input | Mutual Information | The proposed model |
|---|---|---|
| How old are you ? | I’m 4. | I’m 4. why are you asking ? |

# Results

| Input | Mutual Information | The proposed model |
|---|---|---|
| How old are you ? | I’m 4. | I’m 4. why are you asking ? |
| what is your full name ? | i have no idea | what 's yours ? |

# Results

| Input | Mutual Information | The proposed model |
|---|---|---|
| How old are you ? | I’m 4. | I’m 4. why are you asking ? |
| what is your full name ? | i have no idea | what 's yours ? |
| I don 't want to go home tonight . | Really ? | Why ? |

# Results

| Input | Mutual Information | The proposed model |
|---|---|---|
| How old are you ? | I’m 4. | I’m 4. why are you asking ? |
| what is your full name ? | i have no idea | what 's yours ? |
| I don 't want to go home tonight . | Really ? | Why ? |
| Do you have any feelings for me ? | I don’t know what you are talking about. | Would I see you if I didn 't ? |

# Simulation

Keep your expectation low

How old are you ?

# Simulation

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 'm 4, why are you asking ?

I thought you were 12 .

# 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 ?

# 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 '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 .

You don’t know what you are saying.

# 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 .

You don’t know what you are saying.

I don’t know what you are talking about .

# Simulation

Survive 4 turns !!

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 .

You don’t know what you are saying.

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 ?

i 'm 16 .

16 ?

i don 't know what you 're talking about

you don 't know what you 're saying

i don 't know what you 're talking about

you don 't know what you 're saying

# Mutual Information for Response Generation.

Solution 1

$$\hat{T} = \arg\max_{T} \left\{ \log p(T|S) - \lambda \boxed{\log p(T)} \right\}$$

Anti-language Model

Ungrammatical Responses

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

# Mutual Information for Response Generation.

Solution 1 $$\hat{T} = \arg\max_{T} \left\{ \log p(T|S) - \lambda \boxed{\log p(T)} \right\}$$

$$p(T) = \prod_{i=1}^{L_t} p(t_i|t_1, t_2, ..., t_{i-1})$$

# Mutual Information for Response Generation.

Solution 1 $$\hat{T} = \arg\max_{T} \{ \log p(T|S) - \lambda \boxed{\log p(T)} \}$$

$$p(T) = \prod_{i=1}^{L_t} p(t_i|t_1, t_2, ..., t_{i-1})$$

# 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 $r_1, r_2, ..., r_{i-1}$

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

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

# Notations

Action: Generating current response $r_i$

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

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

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

reward $R(r_i)$

# Notations

Action: Generating current response $r_i$

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

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

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

reward $R(r_i)$

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

# Notations

Action: Generating current response $r_i$

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

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

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

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

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

# Notations

Action: Generating current response $r_i$ Infinite #

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

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

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

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

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

# Standard Seq2Seq2