

# Deep Learning in Open Domain Dialogue Generation

Jiwei Li

Computer Science Department

Stanford University

May 10, 2017





- The voice-driven personal assistant on your iPhone
- Perhaps the most visible & exciting application of NLU today
- A major breakthrough in artificial intelligence (AI)??
- The next generation of interaction design??

Borrowed From Bill MacCartney's slides



Does Siri really understand language ?



Does Siri really understand language ?





# Does Siri really understand language ?



**Colbert:** Write the show.



# Does Siri really understand language ?



**Colbert:** Write the show.

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



# Does Siri really understand language ?



**Colbert:** Write the show.

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

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



# Does Siri really understand language ?



**Colbert:** Write the show.

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

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

**Siri:** Searching the Web for "search for anything. I want to write the shuffle."



# Does Siri really understand language ?



**Colbert:** Write the show.

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

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

**Siri:** Searching the Web for "search for anything. I want to write the shuffle."

[...]



# Does Siri really understand language ?



**Colbert:** Write the show.

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

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

**Siri:** Searching the Web for "search for anything. I want to write the shuffle."

[...]

**Colbert:** For the love of God, the cameras are on, give me something?



# Does Siri really understand language ?



**Colbert:** Write the show.

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

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

**Siri:** Searching the Web for "search for anything. I want to write the shuffle."

[...]

**Colbert:** For the love of God, the cameras are on, give me something?

**Siri:** What kind of place are you looking for? Camera stores or churches?

Slide Borrowed From Bill MacCartney

# Reminiscent of Eliza (1966)

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

**ELIZA:** What resemblance do you see

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

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

**User:** You don't argue with me.

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

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

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

# How Eliza works

# Key word list

• • • •

1

If a keyword is identified  
manipulate the input

**Else:**

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

# Background



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

# Background



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

## Goal Oriented Tasks



# Why is building a chatbot hard ?

1. Computers need to **understand** what you ask.

# Why is building a chatbot hard ?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask,

# Why is building a chatbot hard ?

1. Computers need to **understand** what you ask.
2. Computers need to generate coherent, meaningful sequences in response to what you ask, that require **domain knowledge**, **discourse knowledge**, **world knowledge**

# Why is building a chatbot hard ?

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

# Why is building a chatbot hard ?

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

# Background

1. IR based model
2. Generation models

# IR based model

## A big conversation corpus

A: How old are you

B: I am eight

A: What's your name ?

B: I am john

A: How do you like CS224n?

B: I cannot hate it more.

A: How do you like Jiwei ?

B: He's such a Jerk !!!!!

# IR based model

## A big conversation corpus

A: How old are you

B: I am eight

A: What's your name ?

B: I am john

A: How do you like CS224n?

B: I cannot hate it more.

A: How do you like Jiwei ?

B: He's such a Jerk !!!!!

An new input :

What's your age ?

# IR based model

## A big conversation corpus

A: How old are you  
B: I am eight

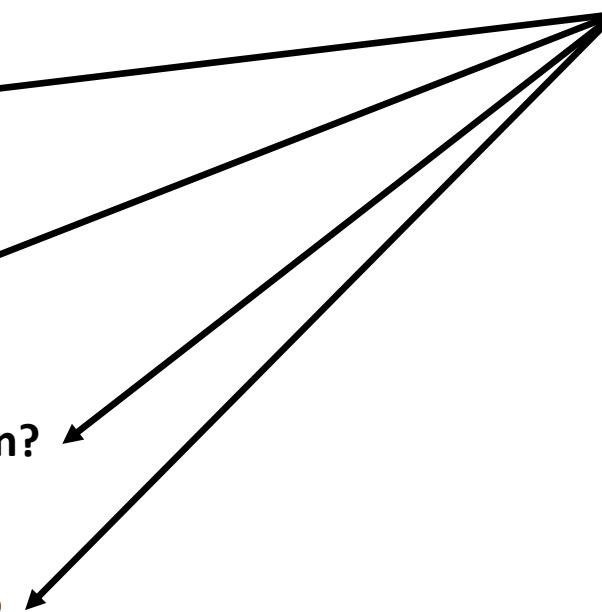
A: What's your name ?  
B: I am john

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

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

An new input :

What's your age ?



# IR based model

## A big conversation corpus

A: How old are you

B: I am eight

A: What's your name ?

B: I am john

A: How do you like CS224n?

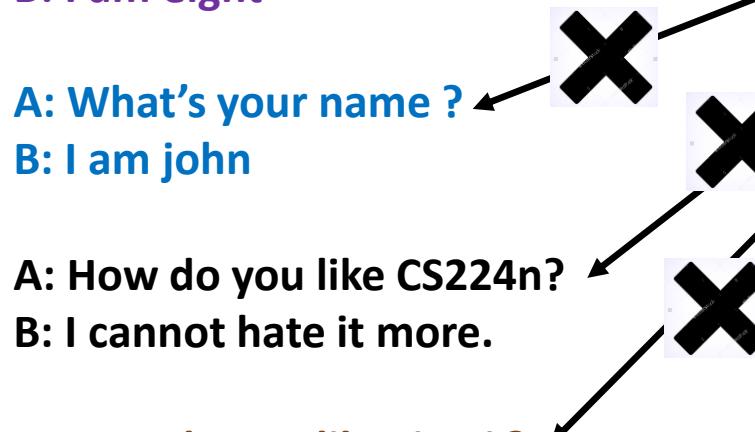
B: I cannot hate it more.

A: How do you like Jiwei ?

B: He's such a Jerk !!!!!

An new input :

What's your age ?



# IR based model

## A big conversation corpus

A: How old are you  
B: I am eight

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

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

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

An new input :

What's your age ?

I am eight.

# Generation Models

# Generation Models

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

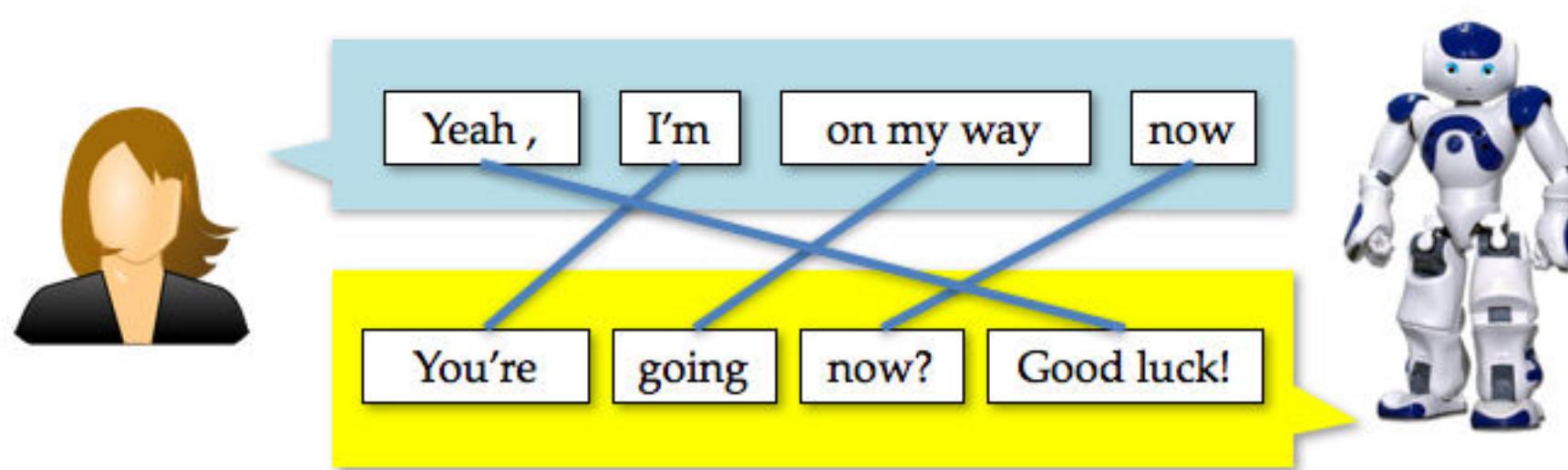


Yeah , I'm on my way now

You're going now? Ok good luck !



# Response Generation as SMT



Exploit high-frequency patterns with phrase-based MT

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

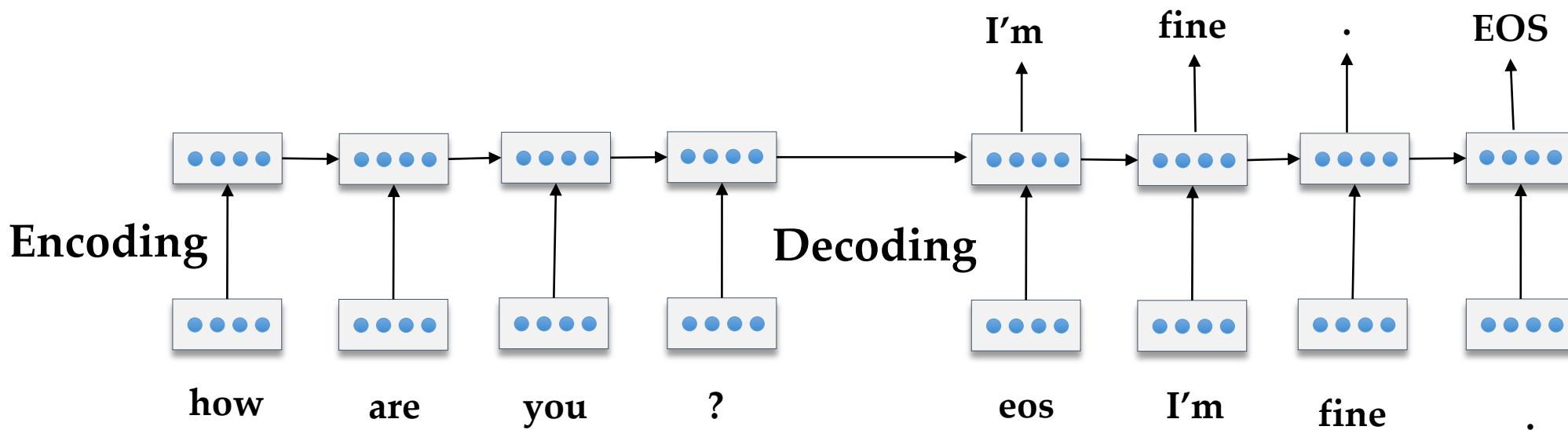
# Seq2Seq Models for Response Generation

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

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

**Source : Input Messages**

**Target : Responses**



# Seq2Seq Models for Response Generation



how



are

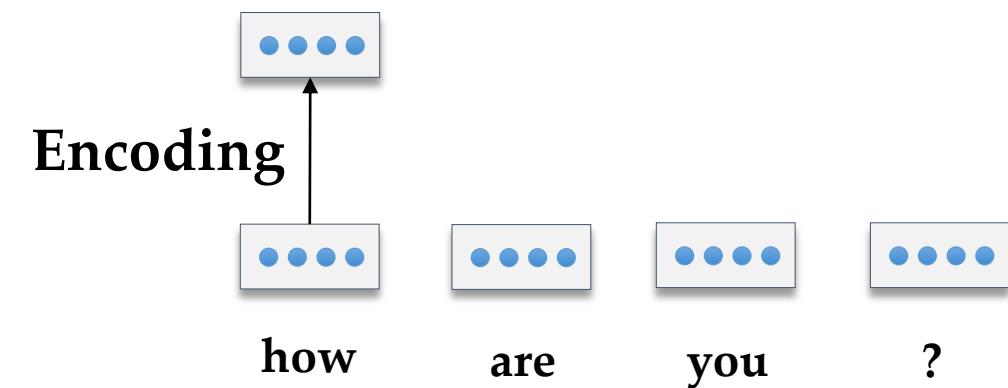


you

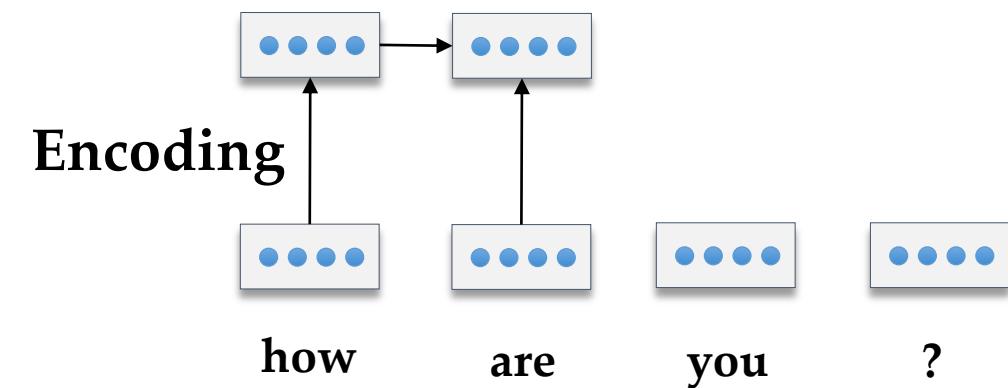


?

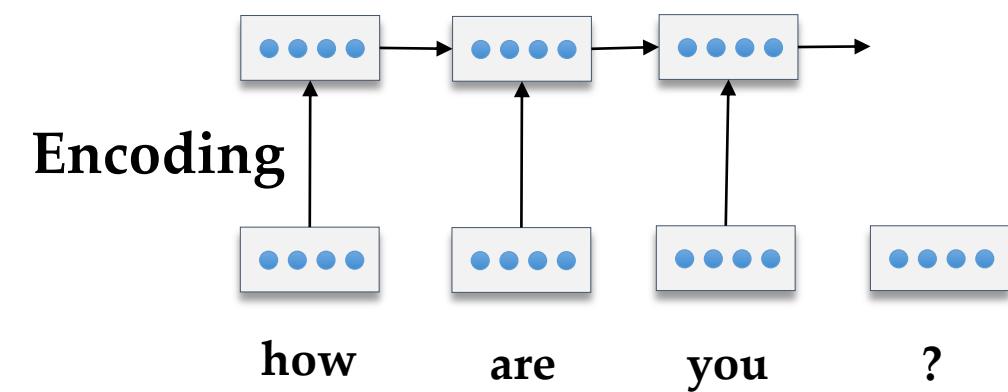
# Seq2Seq Models for Response Generation



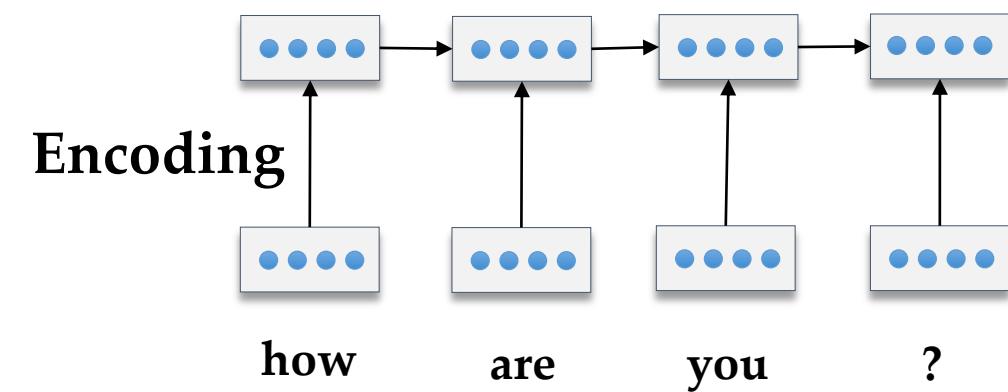
# Seq2Seq Models for Response Generation



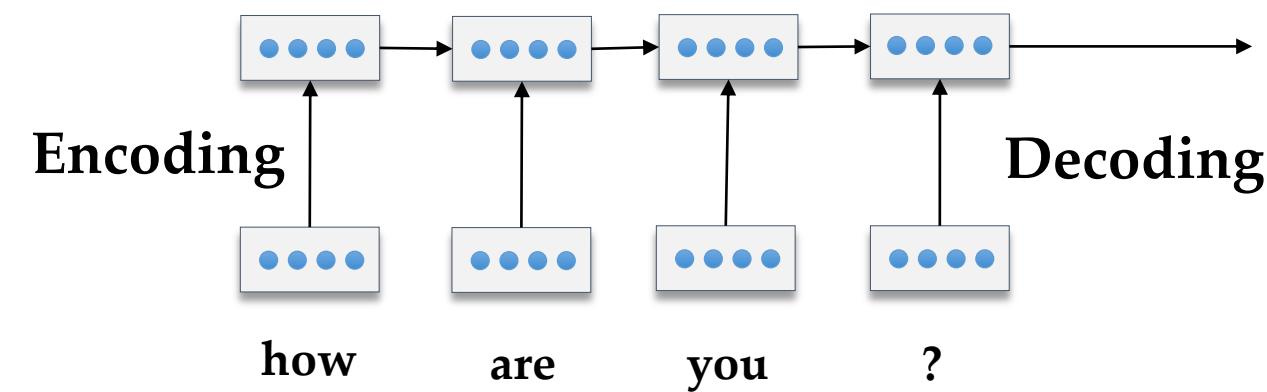
# Seq2Seq Models for Response Generation



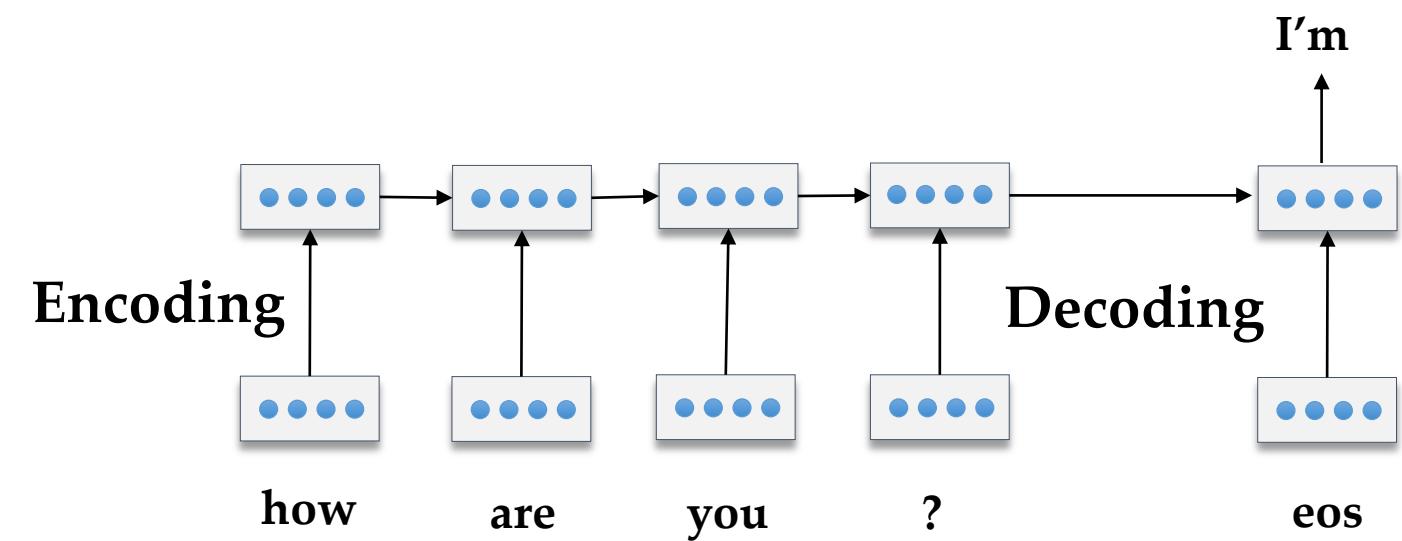
# Seq2Seq Models for Response Generation



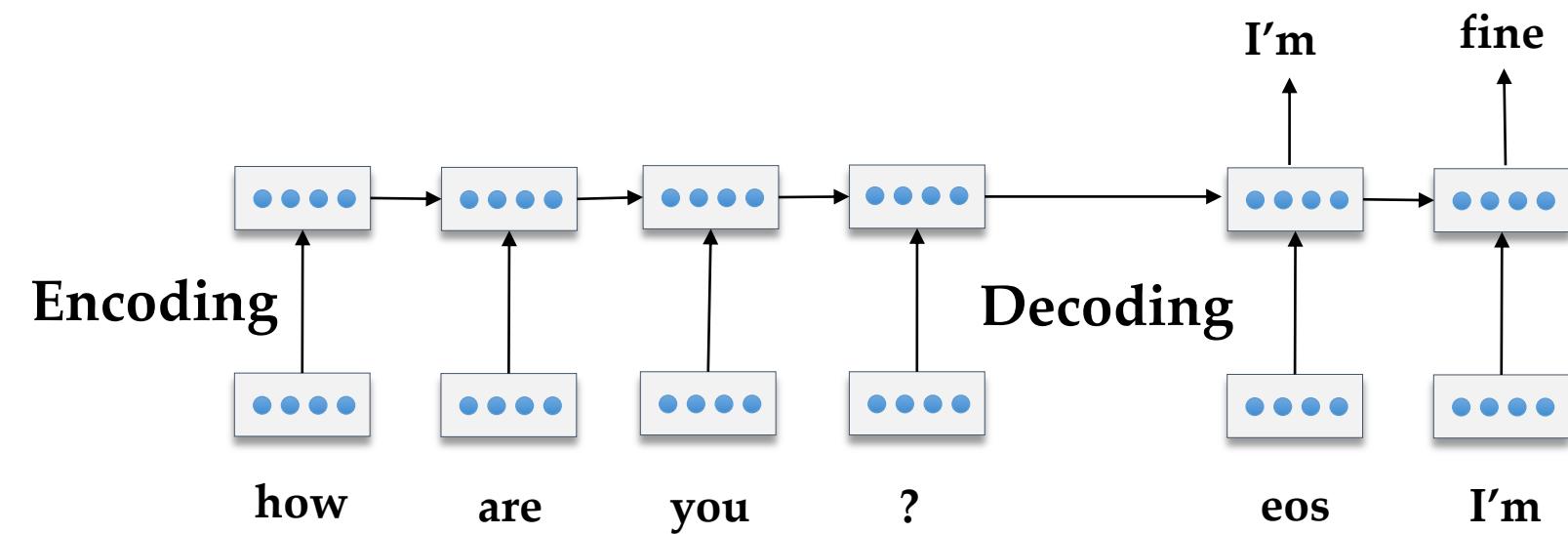
# Seq2Seq Models for Response Generation



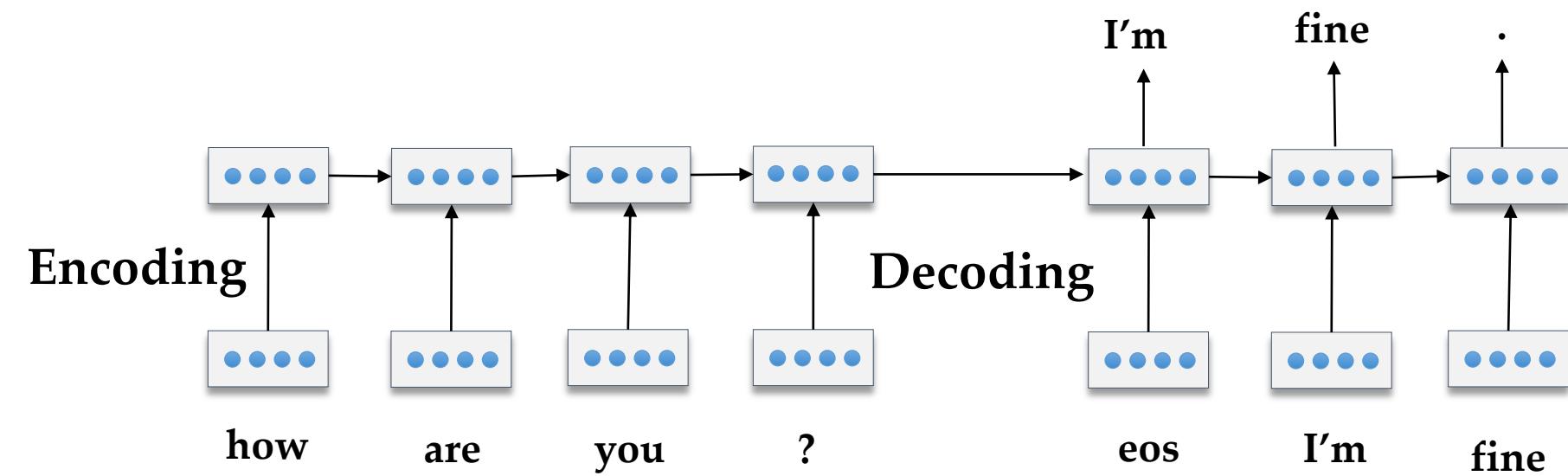
# Seq2Seq Models for Response Generation



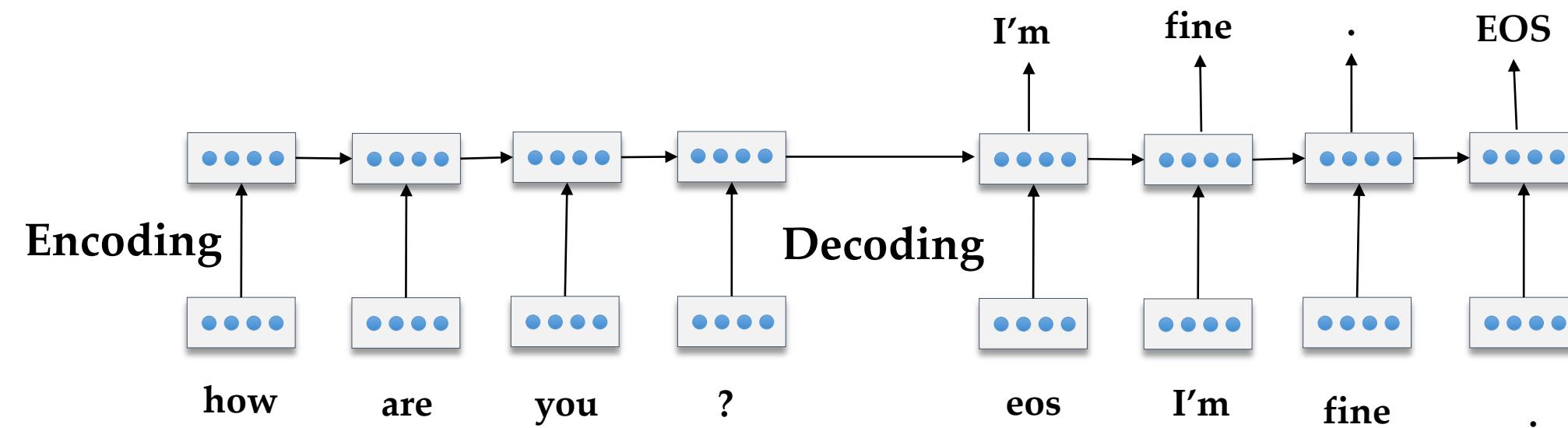
# Seq2Seq Models for Response Generation



# Seq2Seq Models for Response Generation



# Seq2Seq Models for Response Generation



# Sample Results from Google's paper

Input	Response
what is moral ?	what empowered humanity
what is immoral ?	the fact that you have a child .
what is the purpose of existence ?	to find out what happens when we get to the planet earth .
what do you think about bill clinton ?	He's a good man

# Outline

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

# Generic Responses

# Generic Responses

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

# Generic Responses

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



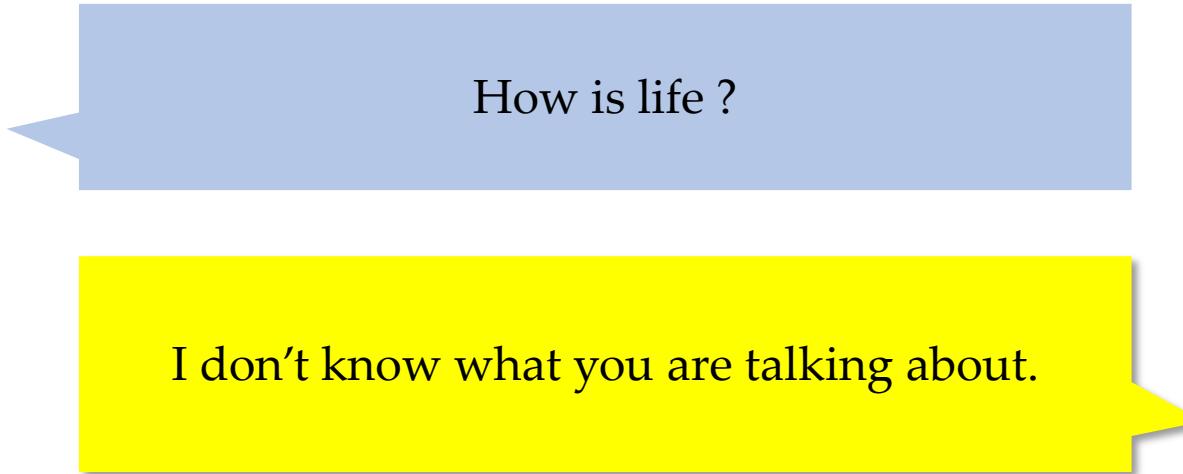
How old are you ?

I don't know .



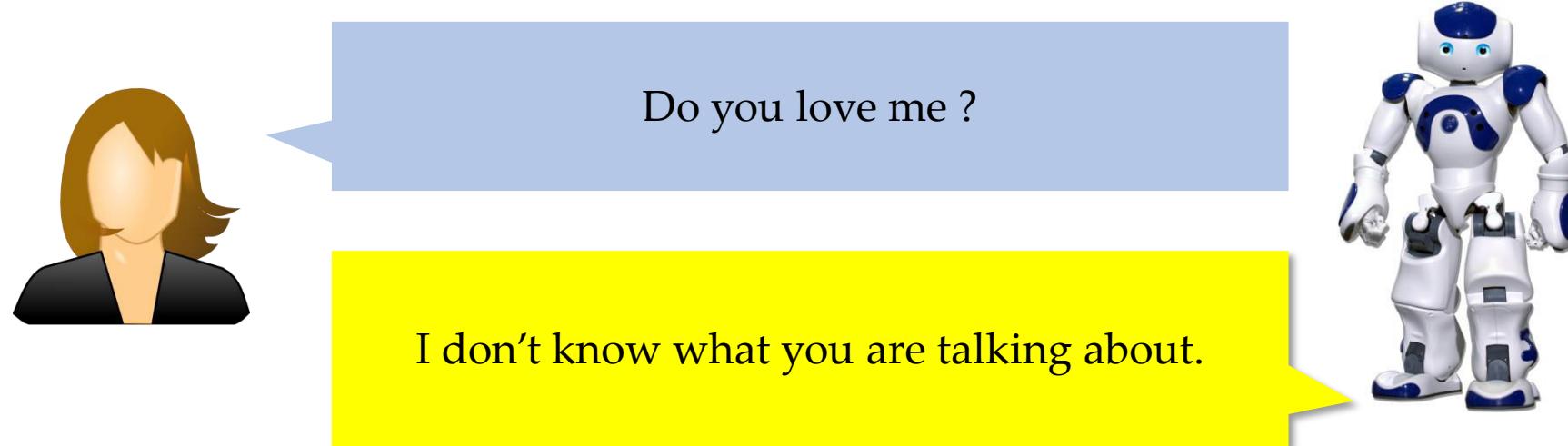
# Generic Responses

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



# How to deal with Generic Responses

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



30% percent of all generated responses

# How to deal with Generic Responses

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

# How to deal with Generic Responses

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

What Eliza is doing !!

# How to deal with Generic Responses

Solution #1: Adding Rules

# How to deal with Generic Responses

## Solution #1: Adding Rules

I don't know .

I don't know ..

I don't know ...

...

I don't know !

I don't know !!

I don't know !!!

# How to deal with Generic Responses

## Solution #1: Adding Rules

I don't know .

I don't know ..

I don't know ...

...

I don't know !

I don't know !!

I don't know !!!

I have no idea .

I don't have a clue.

I 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 ?

# How to deal with Generic Responses

## Solution #1: Adding Rules

I don't know .

I don't know ..

I don't know ...

...

I don't know !

I don't know !!

I don't know !!!

I have no idea .

I don't have a clue.

I 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 ?

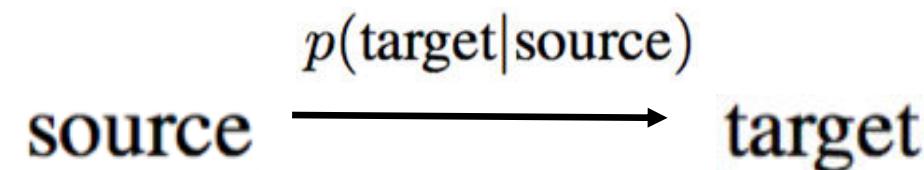
**Rules don't work !!**

# Mutual Information for Response Generation.

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

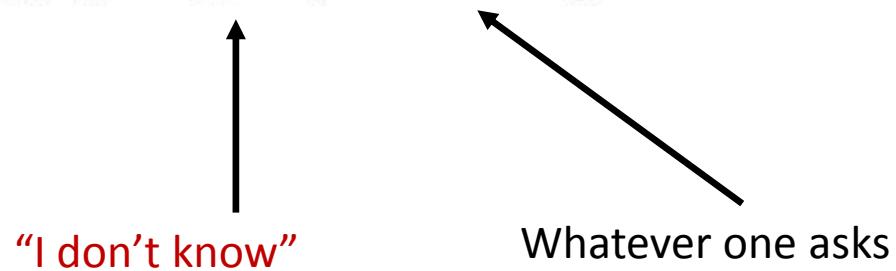
# Mutual Information for Response Generation.

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

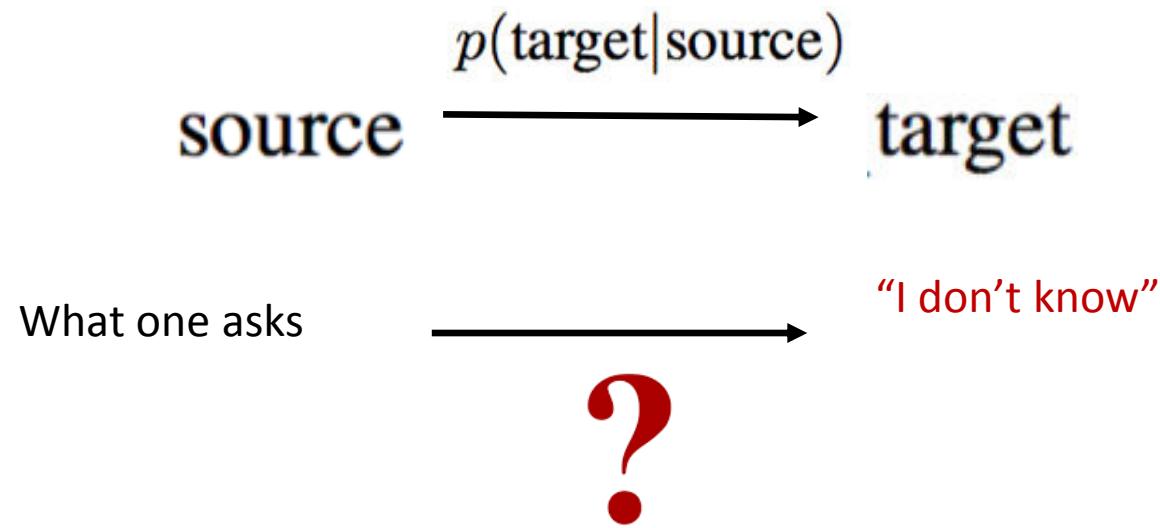


# Mutual Information for Response Generation.

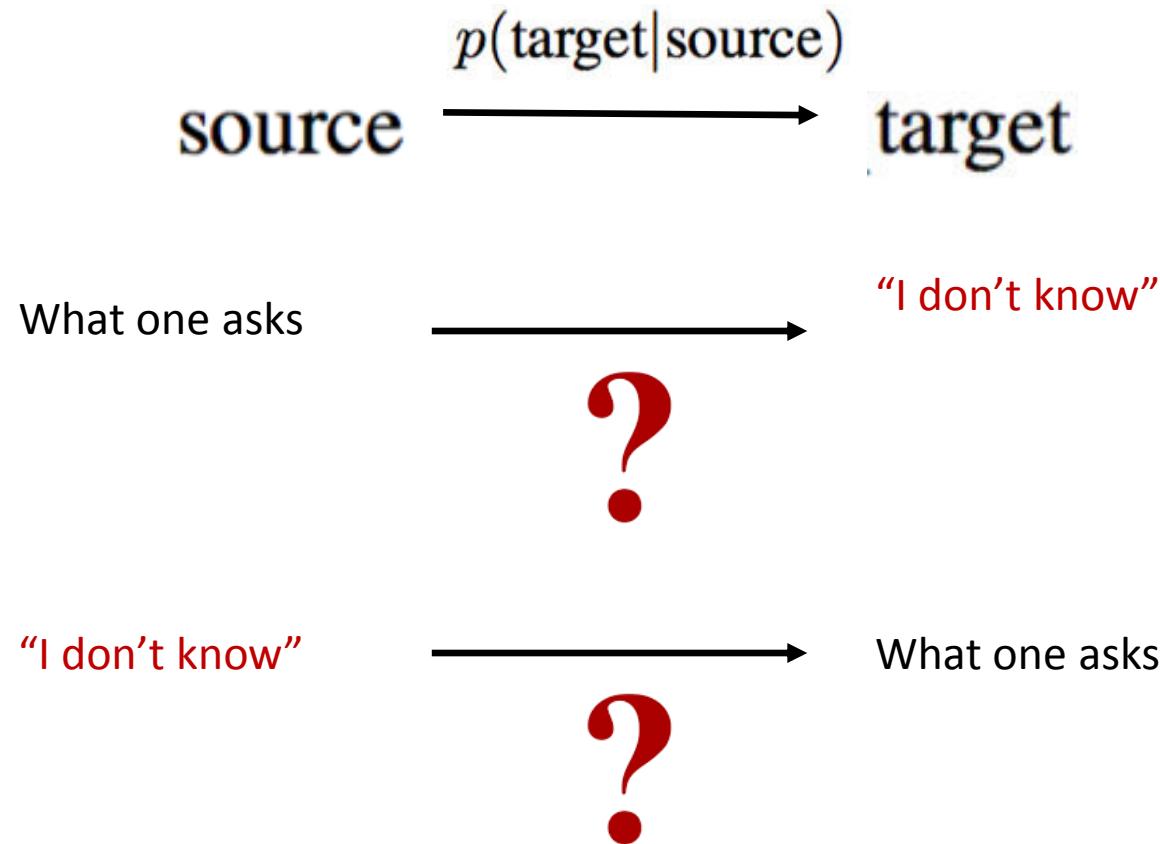
$$\text{Loss} = -\log p(\text{target}|\text{source})$$



# Mutual Information for Response Generation.



# Mutual Information for Response Generation.



# Mutual Information for Response Generation.

$$\begin{array}{ccc} & p(\text{target}|\text{source}) & \\ \text{source} & \xrightarrow{\hspace{2cm}} & \text{target} \\ & p(\text{source}|\text{target}) & \end{array}$$

# Mutual Information for Response Generation.

$$\begin{array}{ccc} & p(\text{target}|\text{source}) & \\ \text{source} & \xrightarrow{\hspace{2cm}} & \text{target} \\ & p(\text{source}|\text{target}) & \end{array}$$

## 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.

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



Bayesian Rule

$$T = \arg \max_T \{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \}$$

# Mutual Information for Response Generation.

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



Bayesian Rule

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

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

# Mutual Information for Response Generation.

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



Bayesian Rule

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

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

# 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

# Outline

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

# Multi-context Response Generation

Single Context:



Any particular plan ?

????

# Multi-context Response Generation



What's your plan for the upcoming summer ?



I am going to Hawaii for 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

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 ?

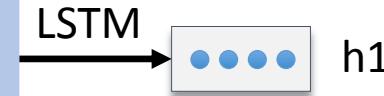
I am going to Hawaii for vocation.

...

Any particular plan ?

# Multi-context Response Generation

What's your plan for the upcoming summer ?



I am going to Hawaii for vocation.



...

...

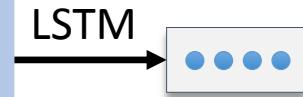


Any particular plan ?



# Multi-context Response Generation

What's your plan for the upcoming summer ?



$c_1$

I am going to Hawaii for vocation.

...



$c_2$

...



$c_K$

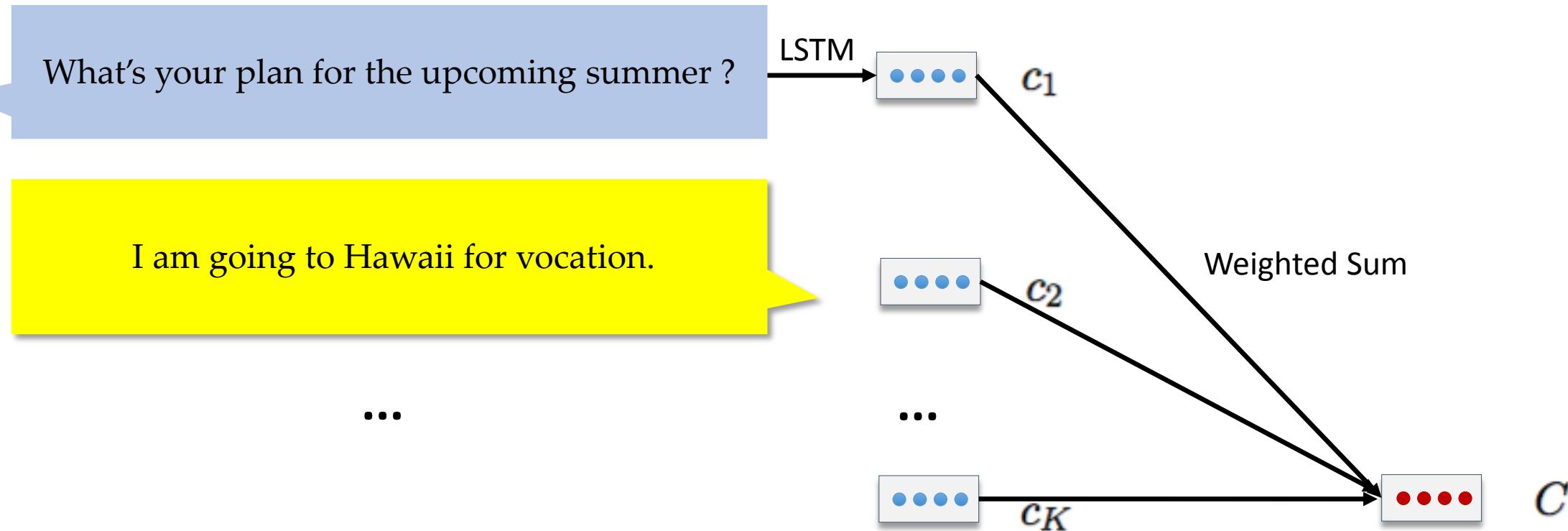
$$c_i = \frac{h_i^T \cdot h_m}{\sum_k \exp(h_i^T h_m)}$$

Any particular plan ?



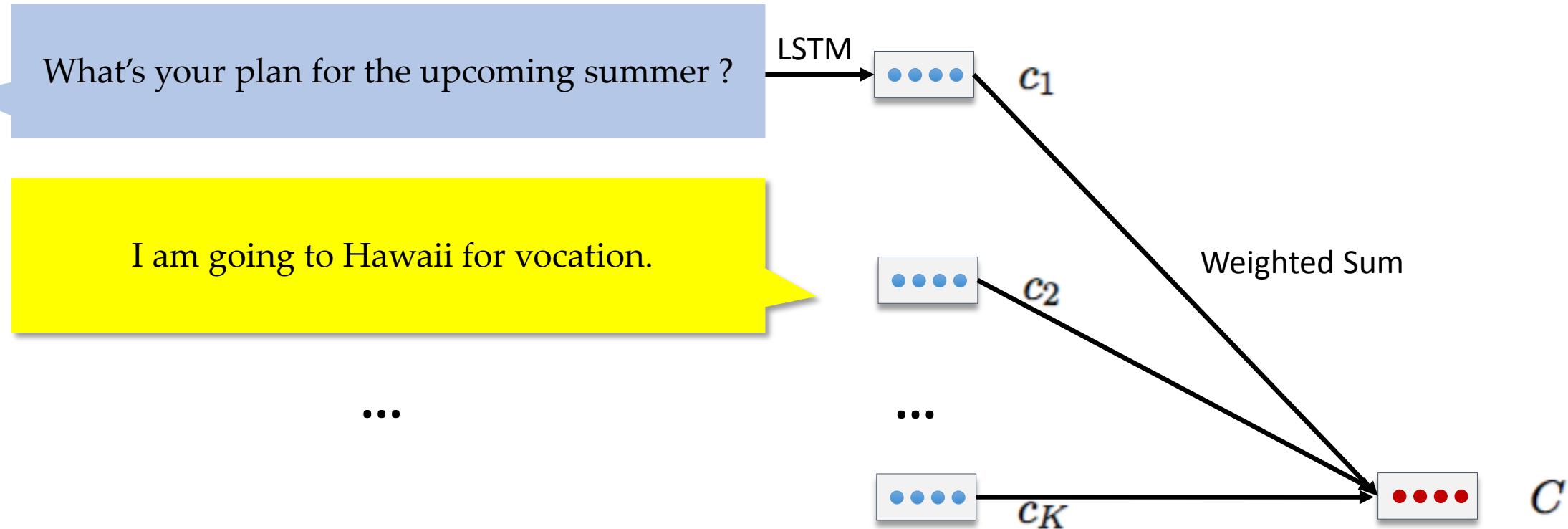
$H$

# Multi-context Response Generation



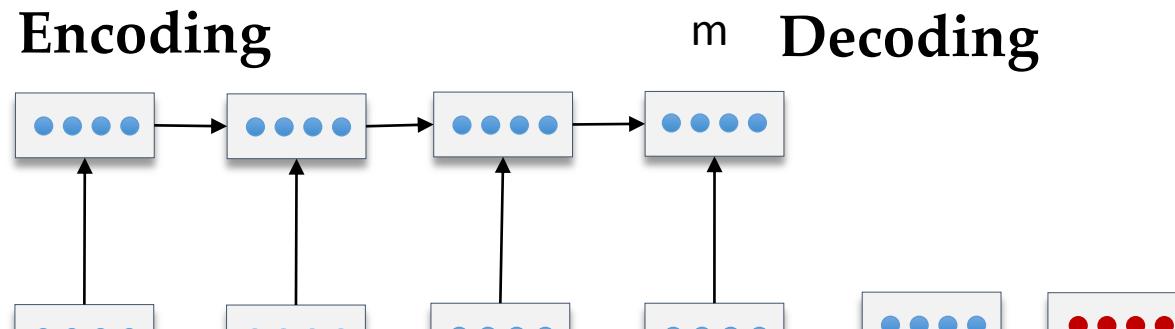
$$C = \sum_i o_i c_i$$

# Multi-context Response Generation



$$C = \sum_i o_i c_i \quad \text{Memory Network (Weston et al., 2014)}$$

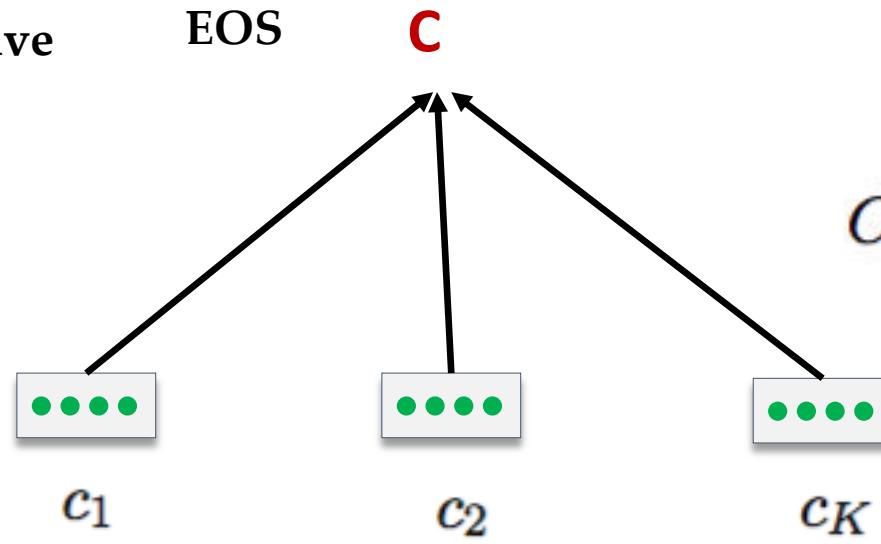
# Multi-context Response Generation



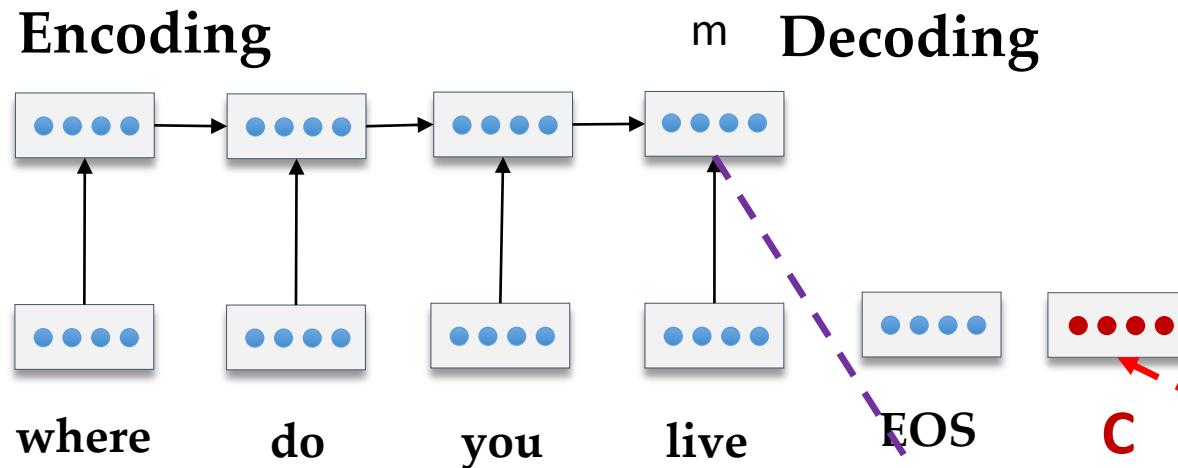
where      do      you      live

EOS

C



# Multi-context Response Generation

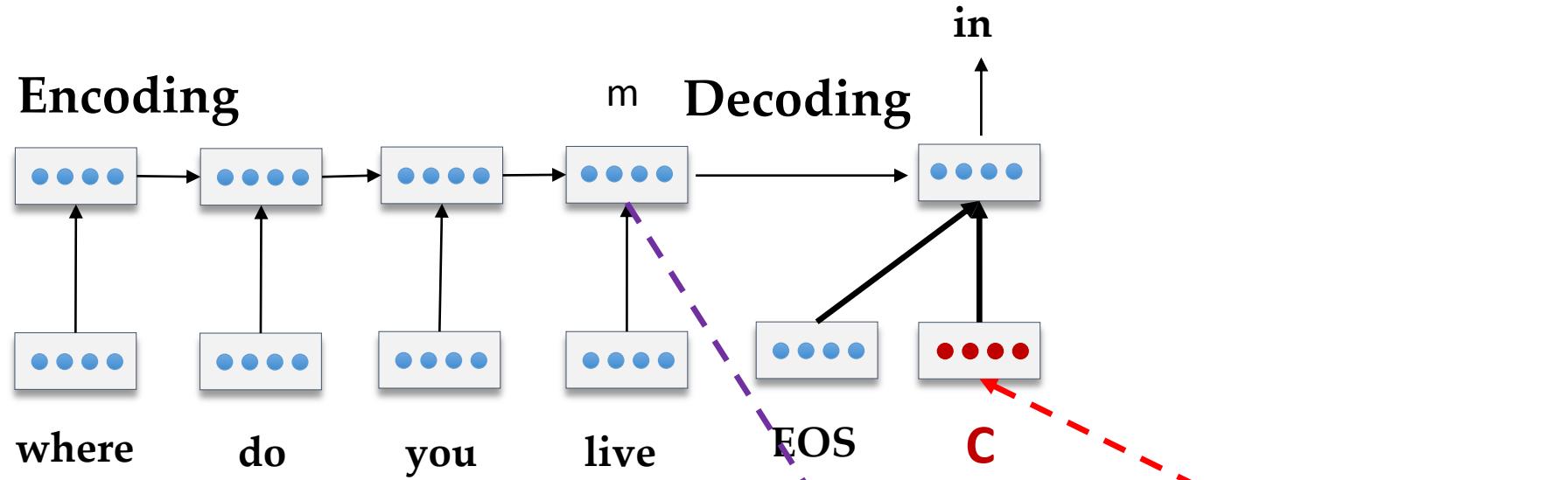


$$O_i = \frac{\exp(c_i^T \cdot h_{t-1})}{\sum_i \exp(c_i^T \cdot h_{t-1})}$$

$$C = \sum_i O_i c_i$$

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

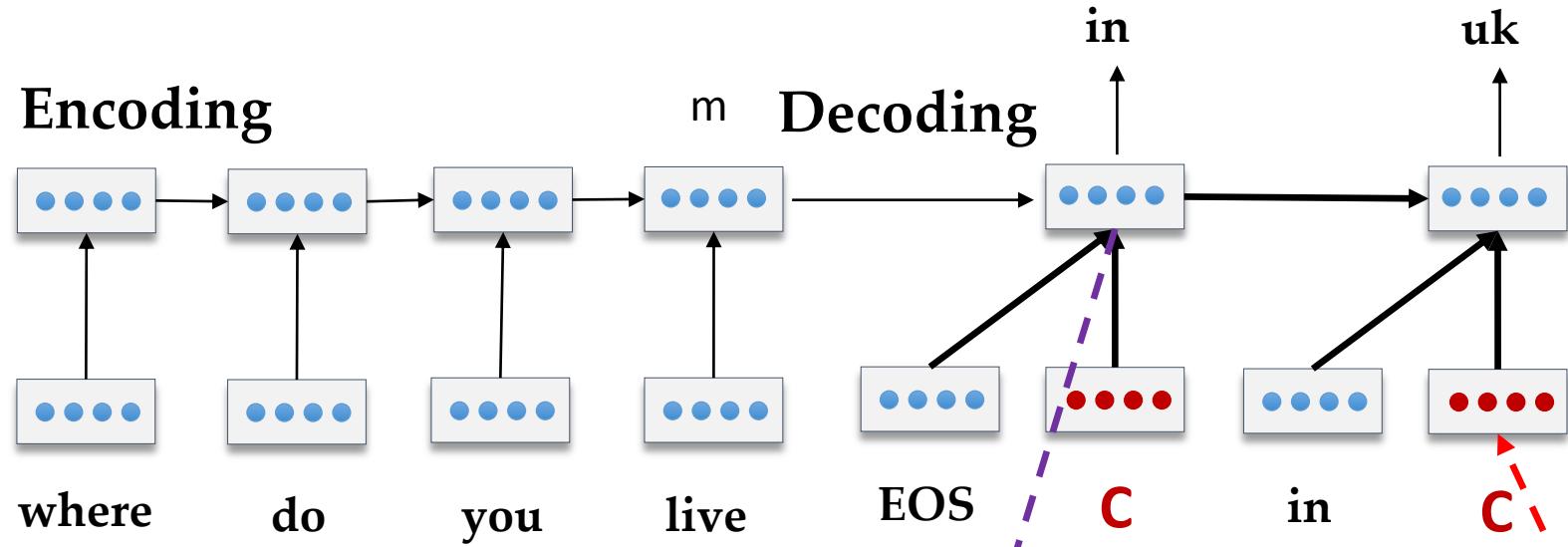
# Multi-context Response Generation



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

$$C = \sum_i O_i c_i$$

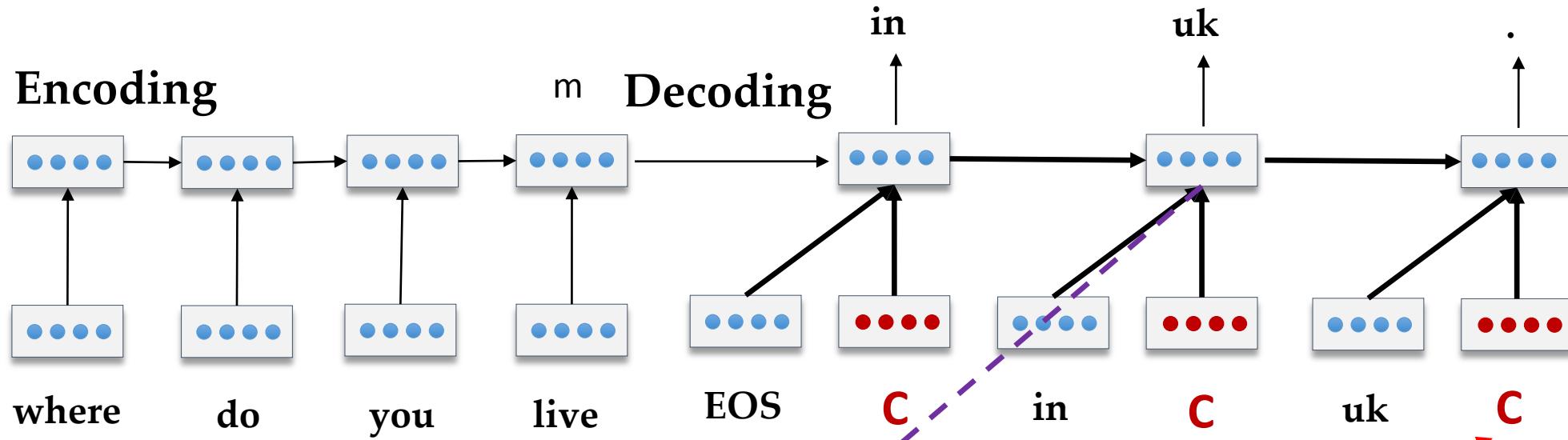
# Multi-context Response Generation



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

$$C = \sum_i O_i c_i$$

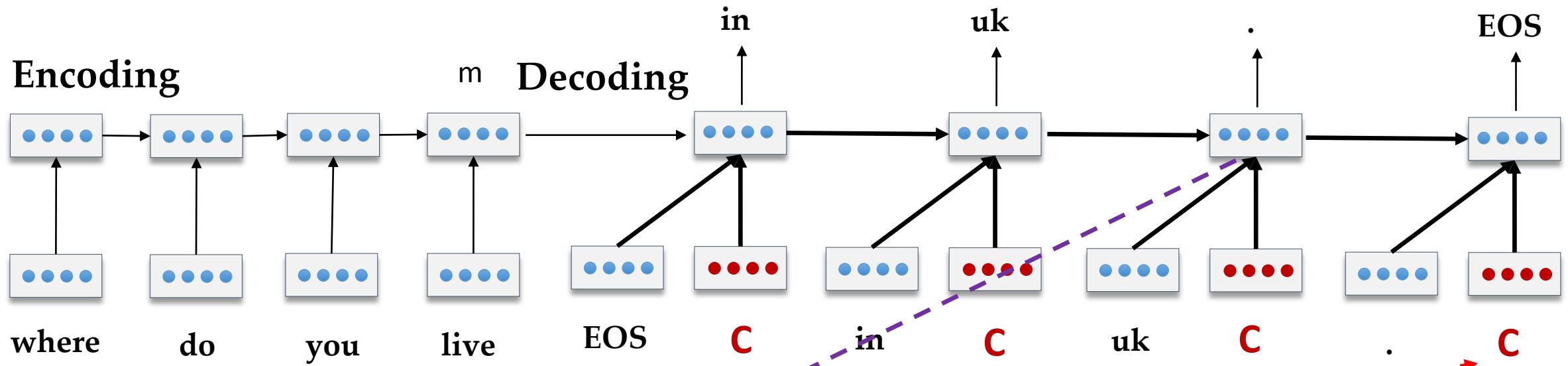
# Multi-context Response Generation



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

$$C = \sum_i O_i c_i$$

# Multi-context Response Generation



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

$$C = \sum_i O_i c_i$$

# Outline

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

# Speaker Consistency

# Speaker Consistency



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 ?

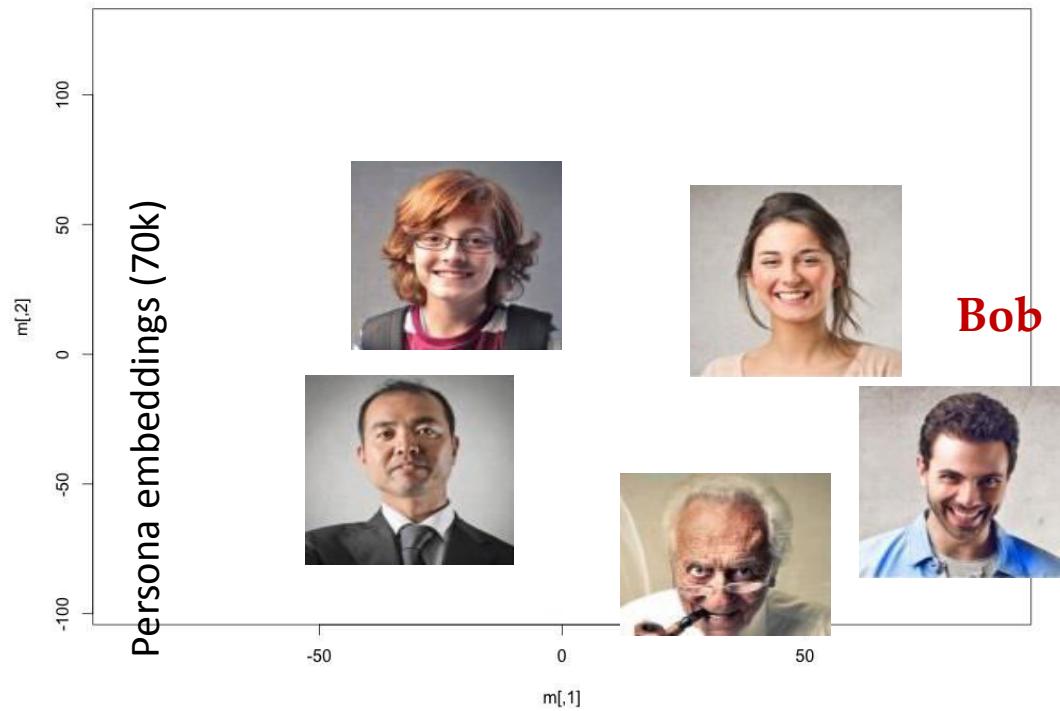
In 1942.



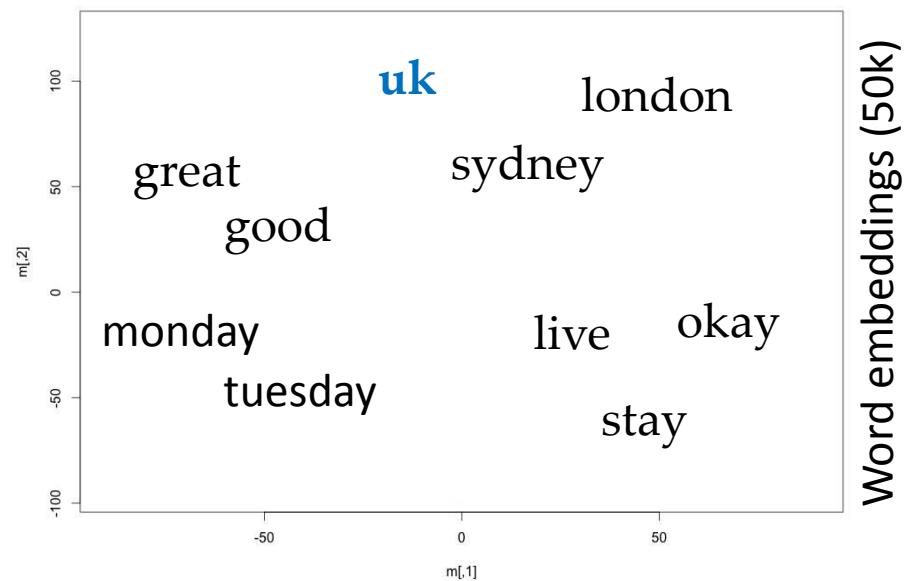
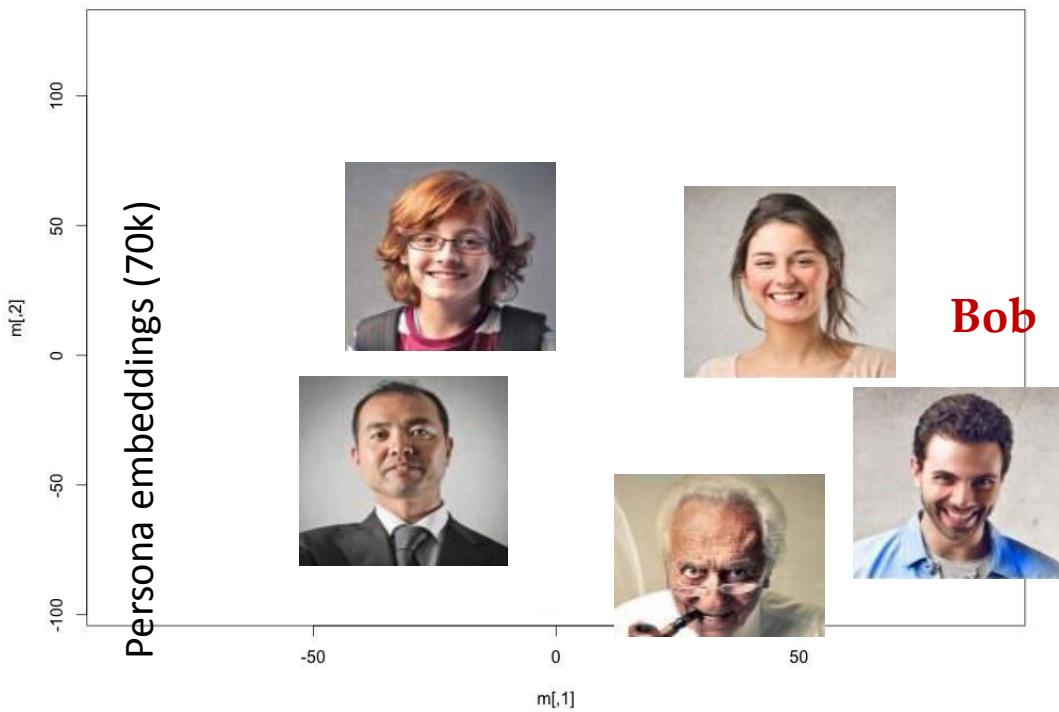
When was your mother born ?

In 1966.

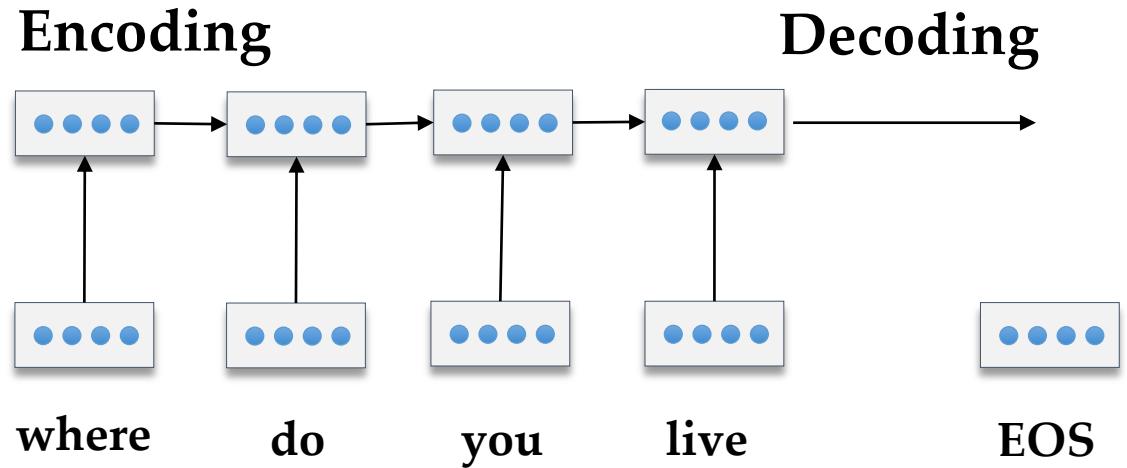
# How to represent users



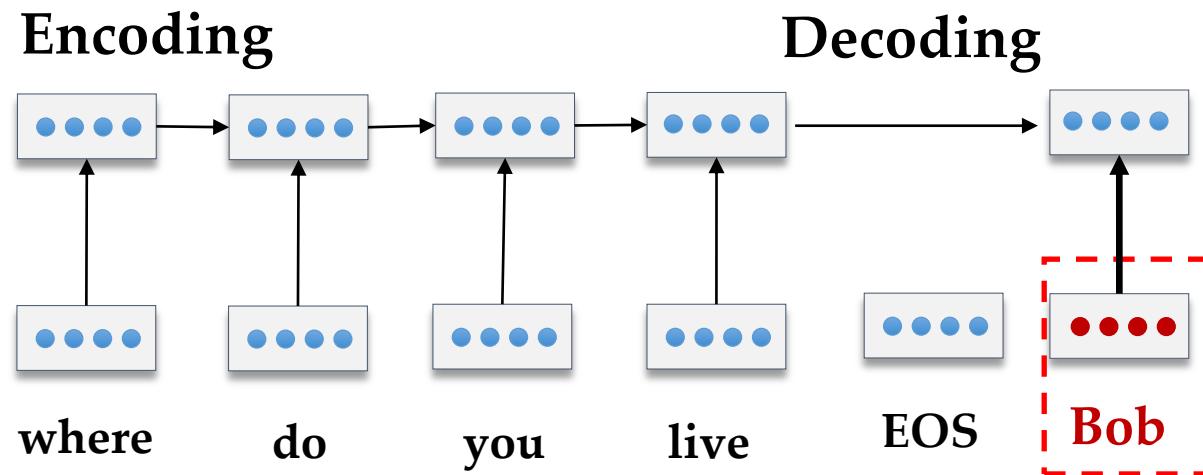
# How to represent users



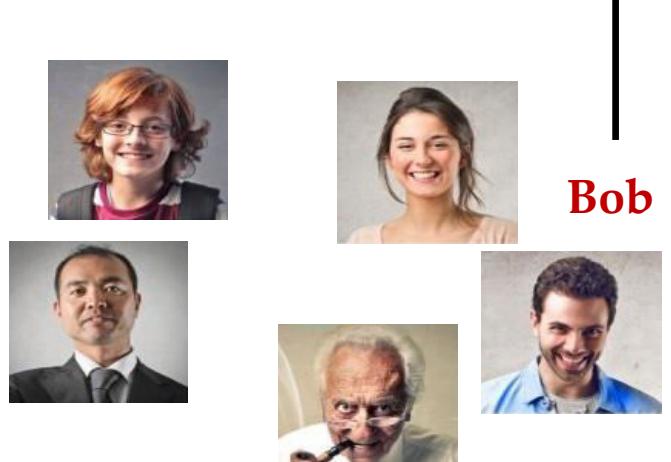
# Persona seq2seq model



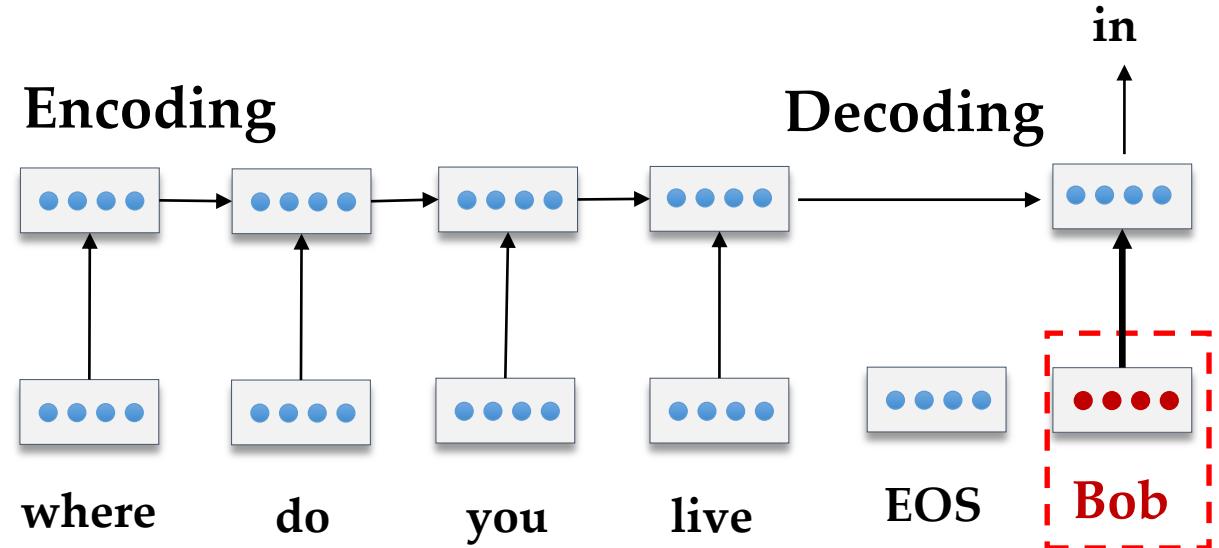
# Persona seq2seq model



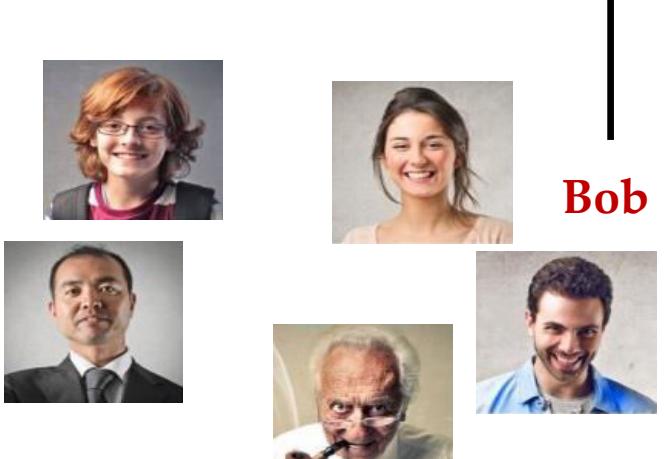
Persona embeddings (70k)



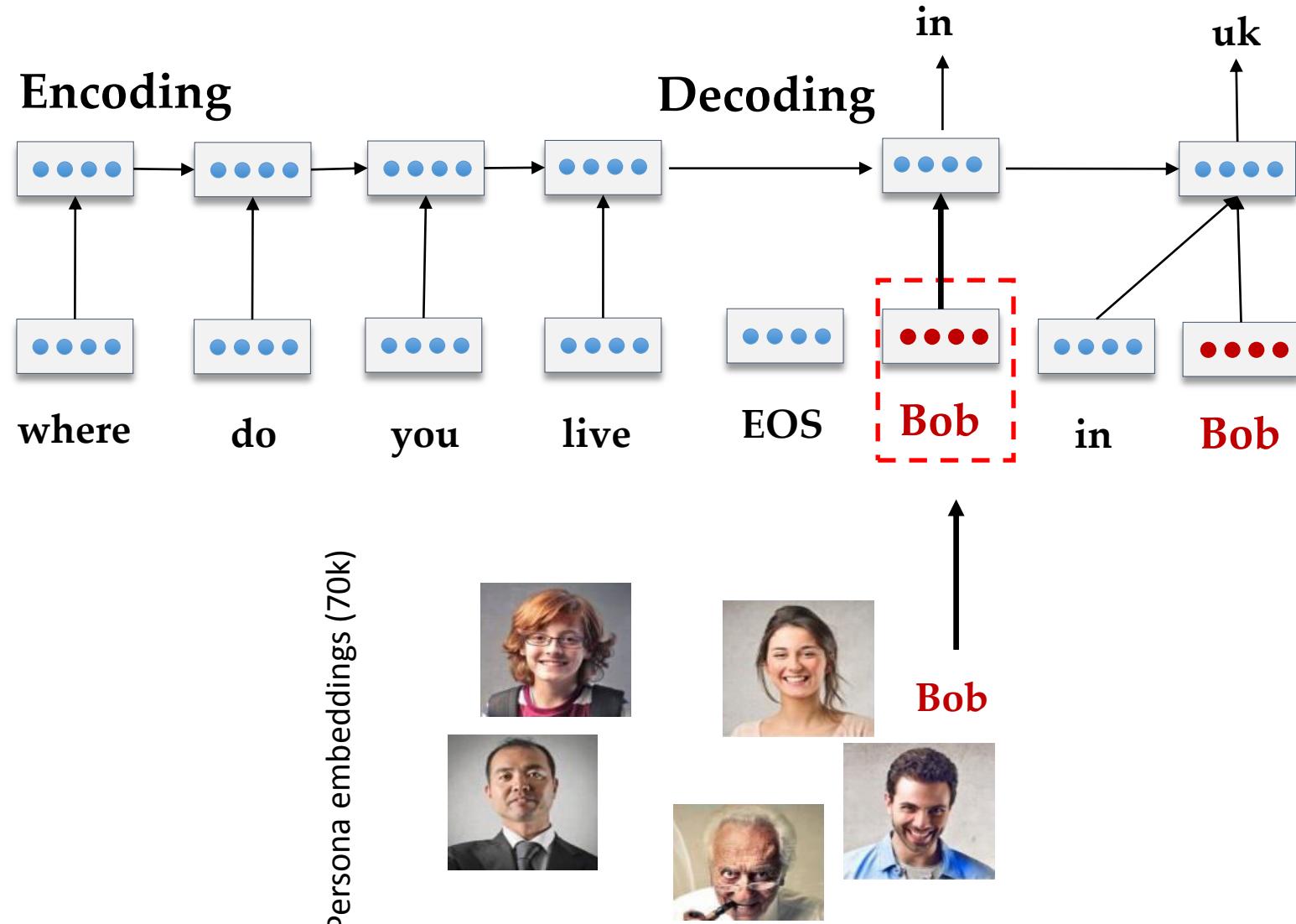
# Persona seq2seq model



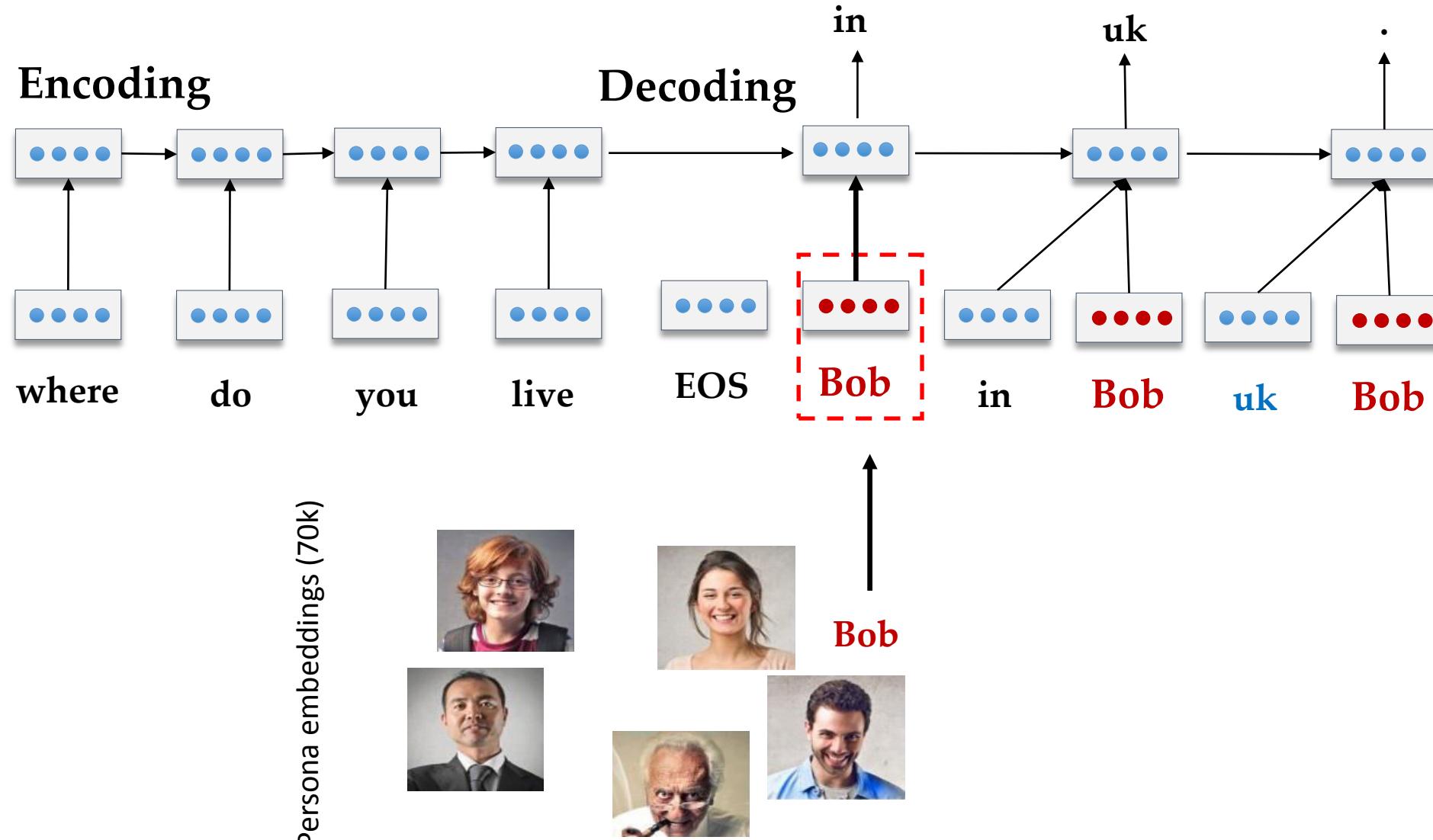
Persona embeddings (70k)



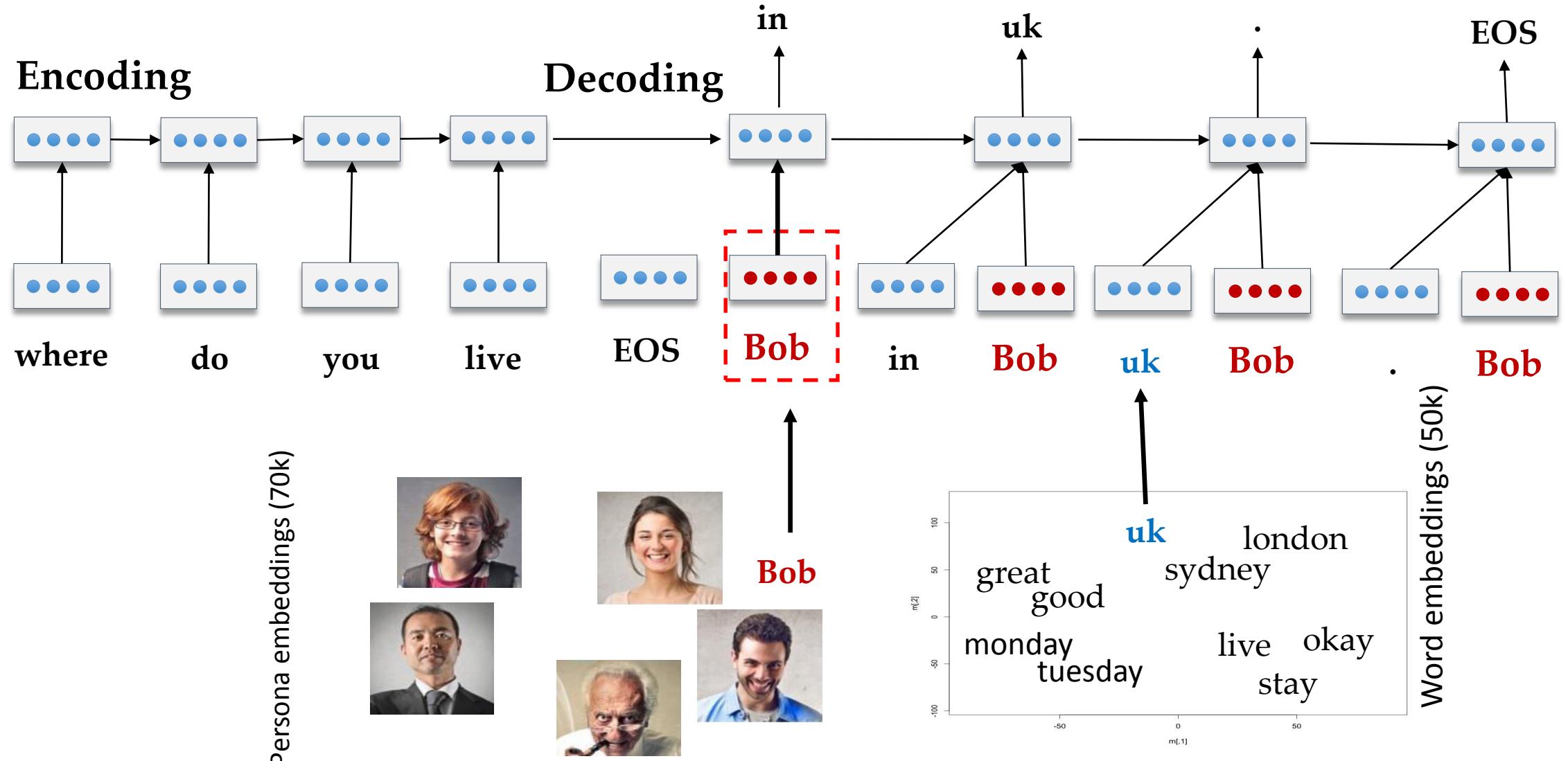
# Persona seq2seq model



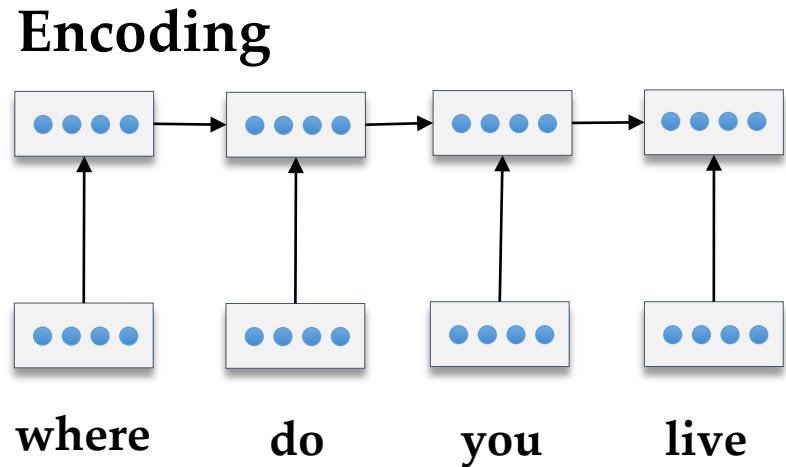
# Persona seq2seq model



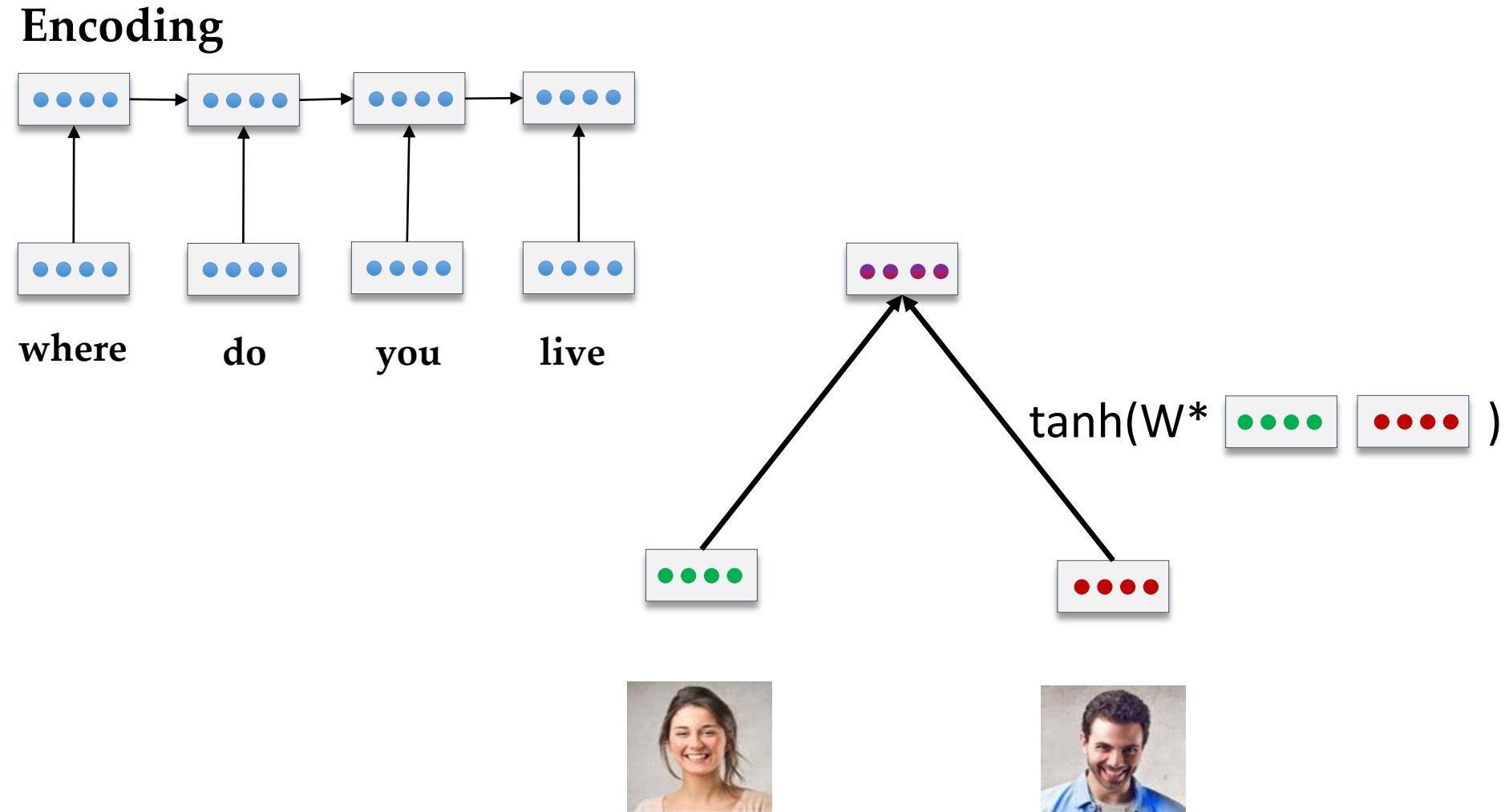
# Persona seq2seq model



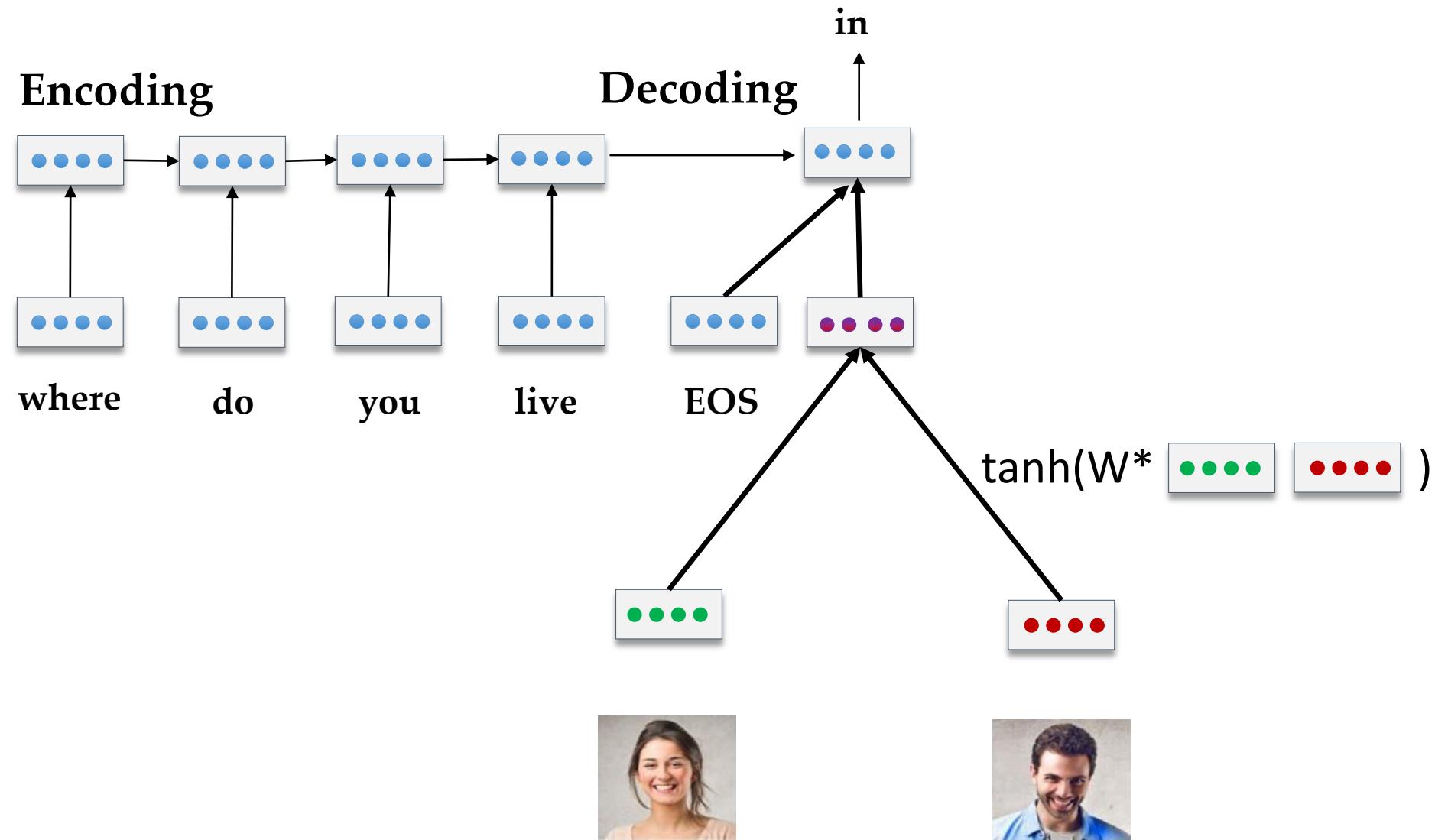
# Interaction Seq2Seq model



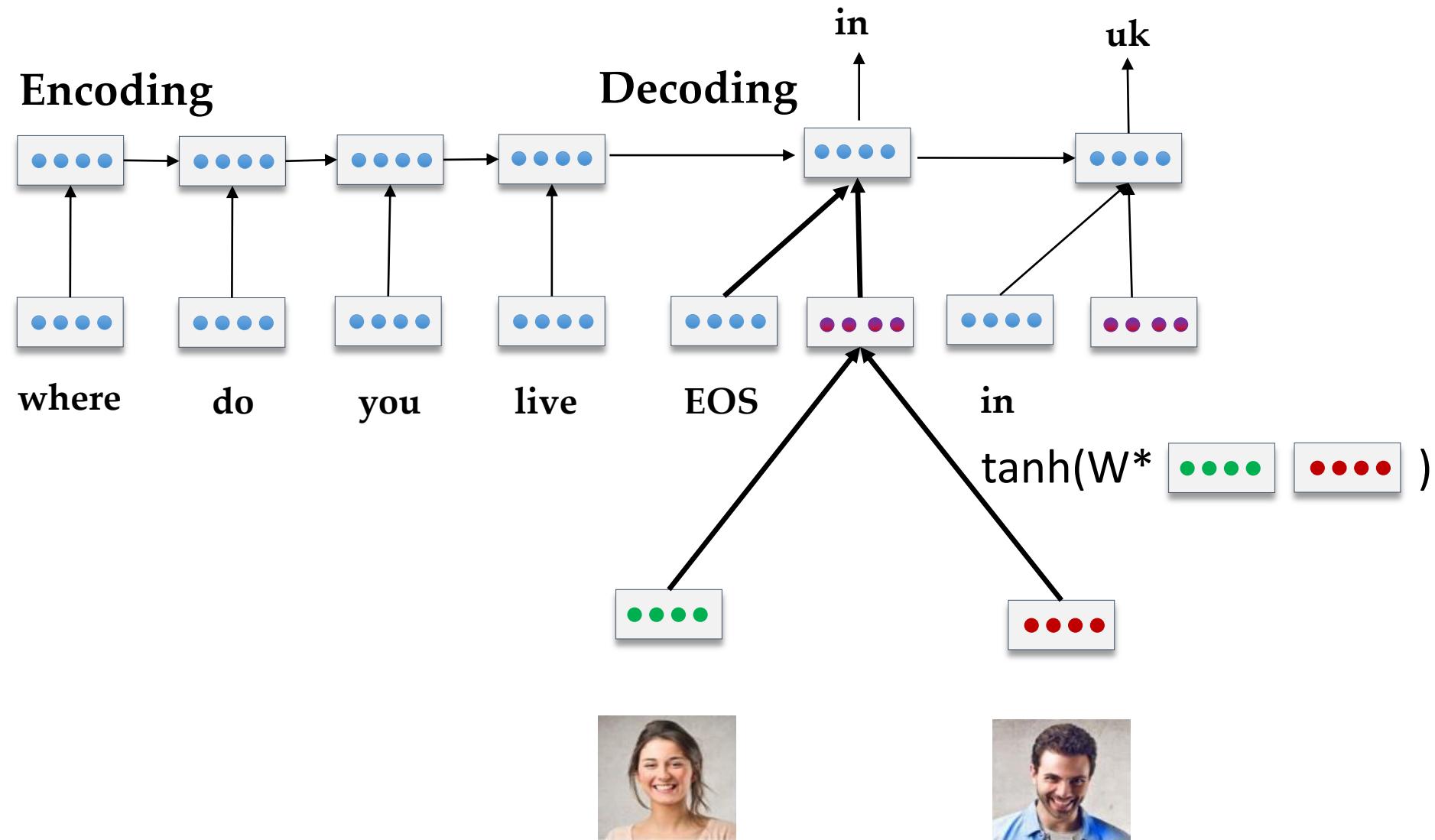
# Interaction Seq2Seq model



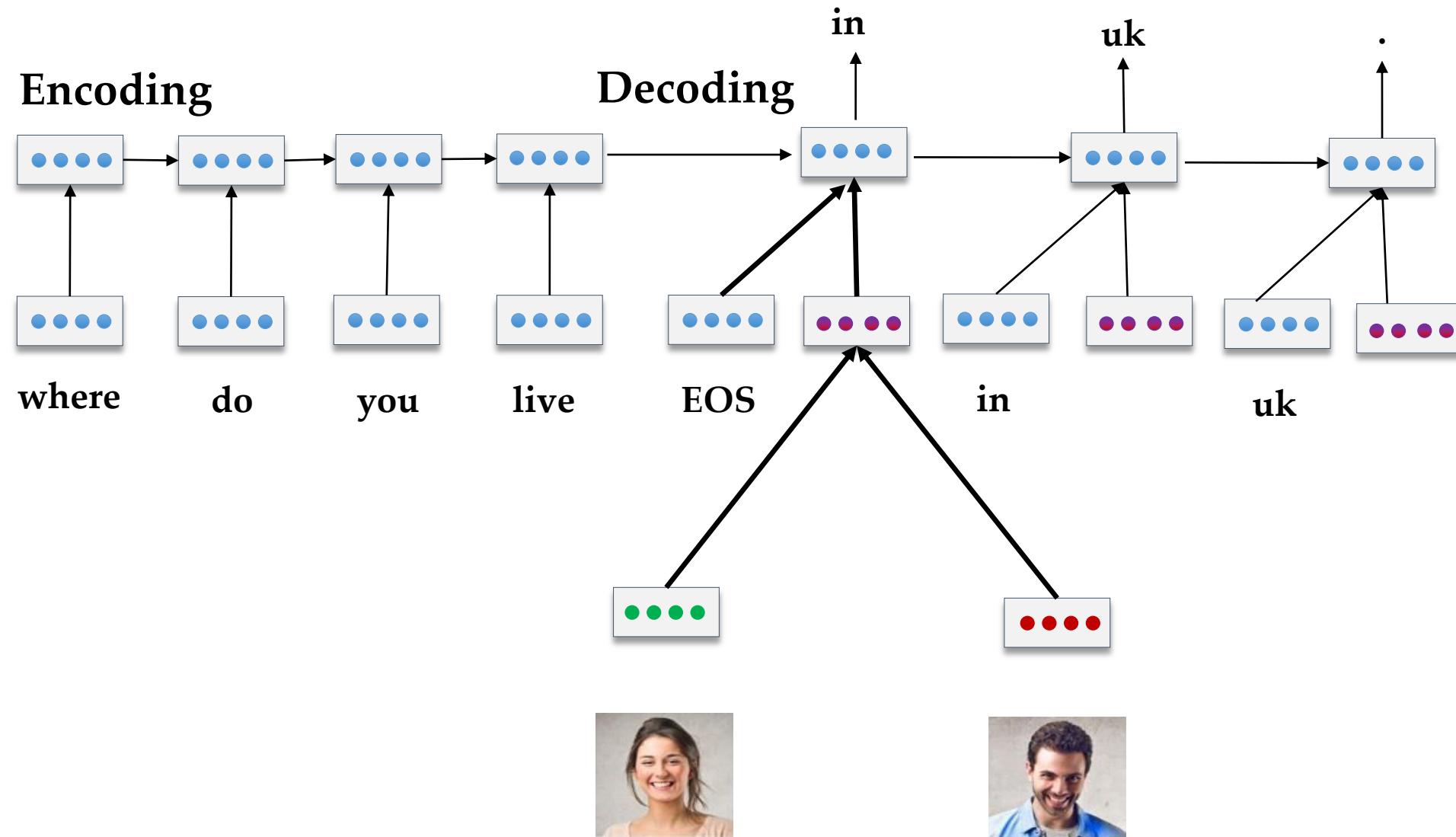
# Interaction Seq2Seq model



# Interaction Seq2Seq model



# Interaction Seq2Seq model



# 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.

---

# 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.

---

# 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.

---

# Results (No cherry-picking)

User1	
<i>message</i>	Where is your hometown?
<i>response</i>	I'm from England.
<i>message</i>	Where are you from?
<i>response</i>	I'm from England.
<i>message</i>	In which city do you live now?
<i>response</i>	I live in London.
<i>message</i>	In which country do you live now?
<i>response</i>	I live in England.
<i>message</i>	What is your major?
<i>response</i>	Business. You?
<i>message</i>	What did you study in college?
<i>response</i>	I did business studies.
<i>message</i>	How old are you?
<i>response</i>	I'm 18.
<i>message</i>	What is your age?
<i>response</i>	I'm 18.

User2	
<i>message</i>	Where is your hometown?
<i>response</i>	I live in Indonesia.
<i>message</i>	Where are you from?
<i>response</i>	England, you?
<i>message</i>	In which city do you live now?
<i>response</i>	I live in Jakarta.
<i>message</i>	In which country do you live now?
<i>response</i>	I live in Indonesia.
<i>message</i>	What is your major?
<i>response</i>	Business, you?
<i>message</i>	What did you study in college?
<i>response</i>	Psychology, you?
<i>message</i>	How old are you?
<i>response</i>	I'm 18.
<i>message</i>	What is your age?
<i>response</i>	I'm 16.

# Issues

How do we handle long-term dialogue success?

# Outline

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

# Issues

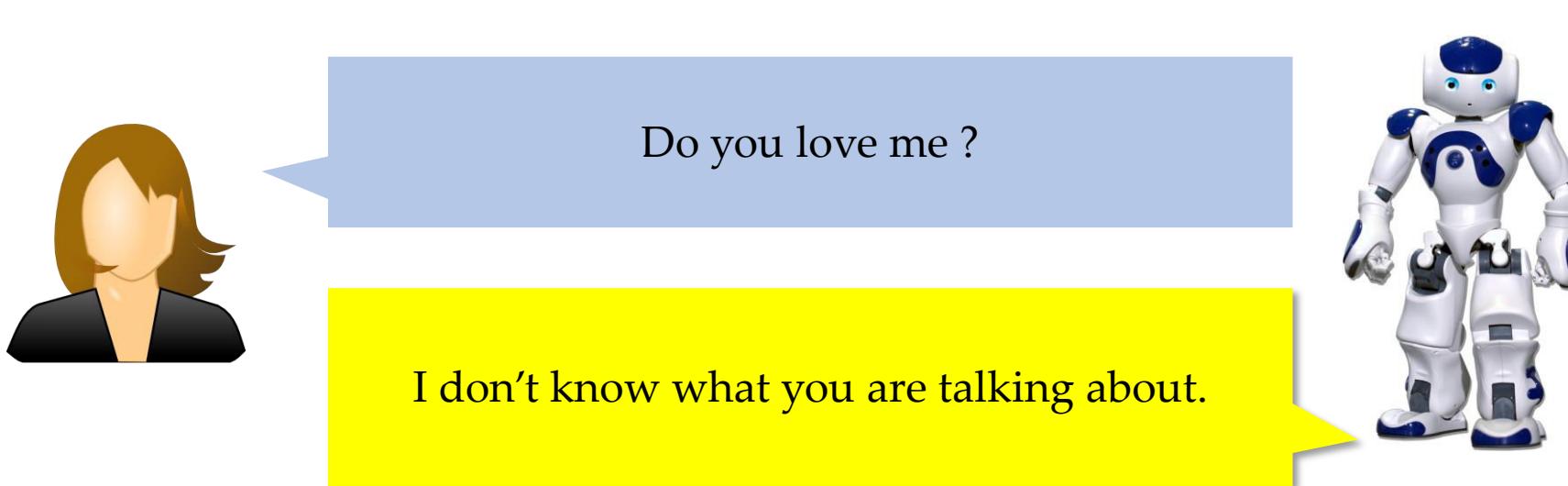
How do we handle long-term dialogue success?

- Problem 1: Dull and generic responses.

# Issues

- Problem 1: Dull and generic responses.

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



# Issues

How do we handle long-term dialogue success?

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

# Problem 2: Repetitive responses.

Shut up !



## Problem 2: Repetitive responses.

Shut up !

No, you shut up !



## Problem 2: Repetitive responses.

Shut up !

No, you shut up !



No, you shut up !



## Problem 2: Repetitive responses.



Shut up !

No, you shut up !

No, you shut up !

No, you shut up !



## Problem 2: Repetitive responses.



Shut up !

No, you shut up !

No, you shut up !

• • • • •

No, you shut up !

No, you shut up !



# Problem 2: Repetitive responses.



See you later !

See you later !

See you later !

• • • • •

See you later !

See you later !



# Issues

How do we handle long-term dialogue success?

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

# Problem 3: Short-sighted conversation decisions.



How old are you ?



# Problem 3: Short-sighted conversation decisions.



How old are you ?

i 'm 16 .



# Problem 3: Short-sighted conversation decisions.



How old are you ?

i 'm 16 .

16 ?



# Problem 3: Short-sighted conversation decisions.



How old are you ?

16 ?

i 'm 16 .

i don 't know what you 're talking about



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .

16 ?

i don 't know what you 're talking about



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don 't know what you 're talking about

you don 't know what you 're saying



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don 't know what you 're talking about

you don 't know what you 're saying

i don 't know what you 're talking about



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don 't know what you 're talking about

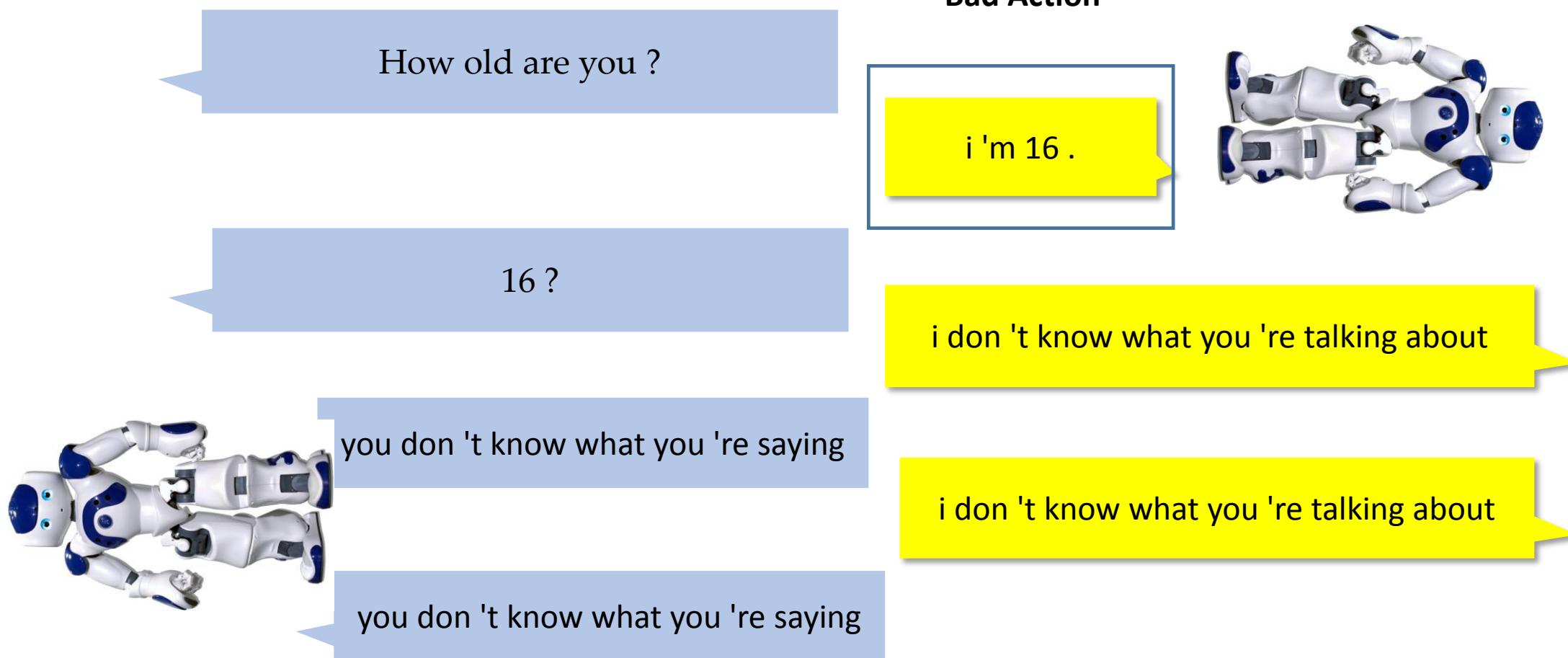
you don 't know what you 're saying

i don 't know what you 're talking about



you don 't know what you 're saying

# Problem 3: Short-sighted conversation decisions.



# Problem 3: Short-sighted conversation decisions.

How old are you ?

i 'm 16 .



16 ?

i don 't know what you 're talking about

you don 't know what you 're saying

i don 't know what you 're talking about

you don 't know what you 're saying



Outcome

# Can reinforcement learning handle this?

How old are you ?

i 'm 16 .



16 ?

i don 't know what you 're talking about

you don 't know what you 're saying

i don 't know what you 're talking about

you don 't know what you 're saying

Outcome does not emerge  
until a few turns later



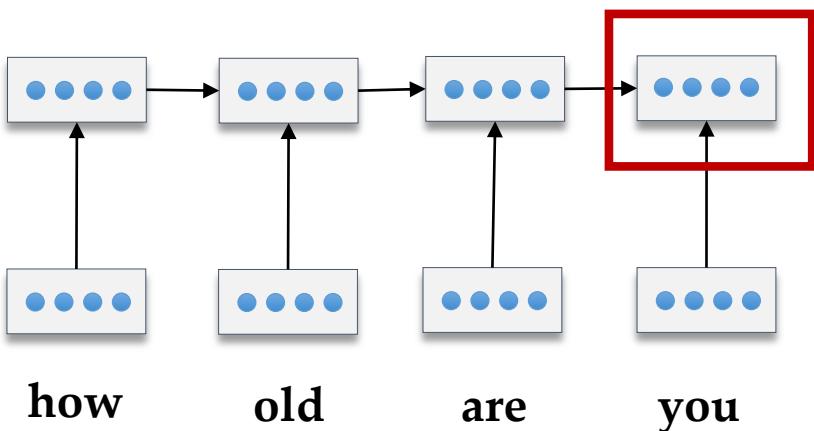
Can reinforcement learning handle this?

# Notations for Reinforcement Learning

# Notations: 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

# Notations: Reward

$r_{i-1}$

How old are you ?



$r_i$

i'm 16 .



# Notations: Reward

1. Ease of answering

# Notations: Reward

1. Ease of answering

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

# Notations: Reward

## 1. Ease of answering

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

S: "I don't know what you are talking about"

# Notations: Reward

## 2. Information Flow

# Notations: Reward

## 2. Information Flow



See you later !



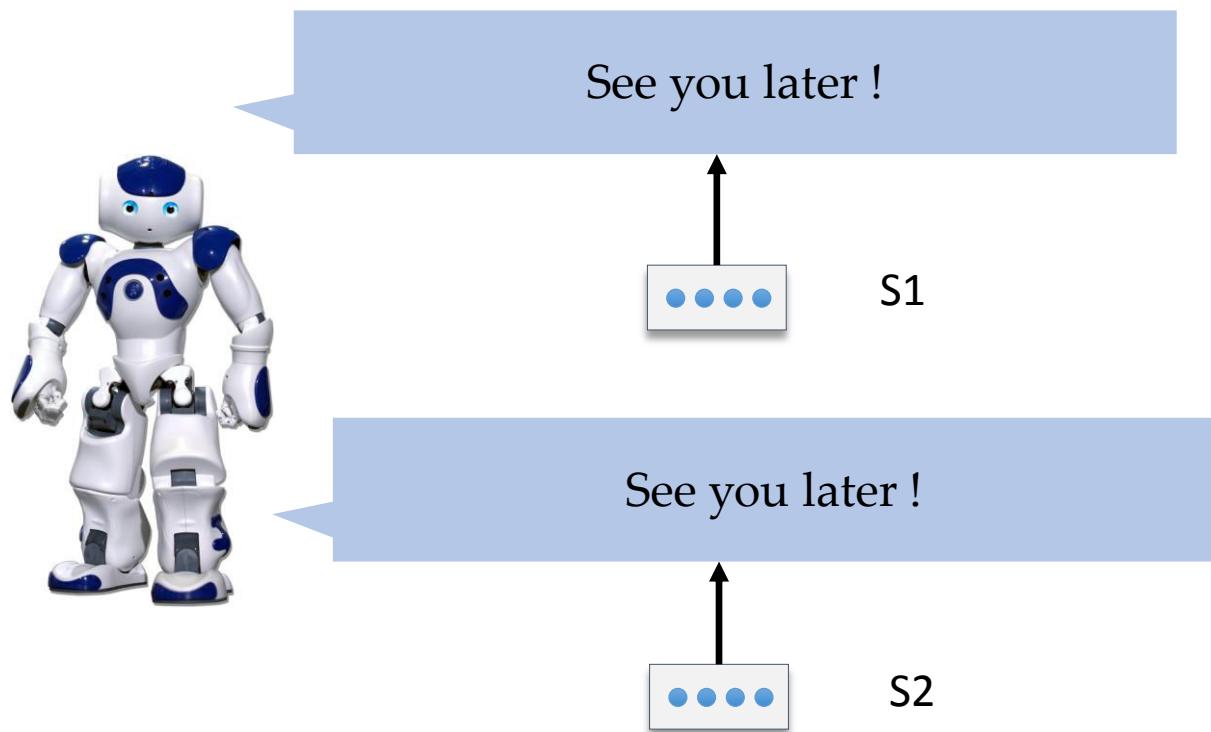
See you later !

See you later !

See you later !

# Notations: Reward

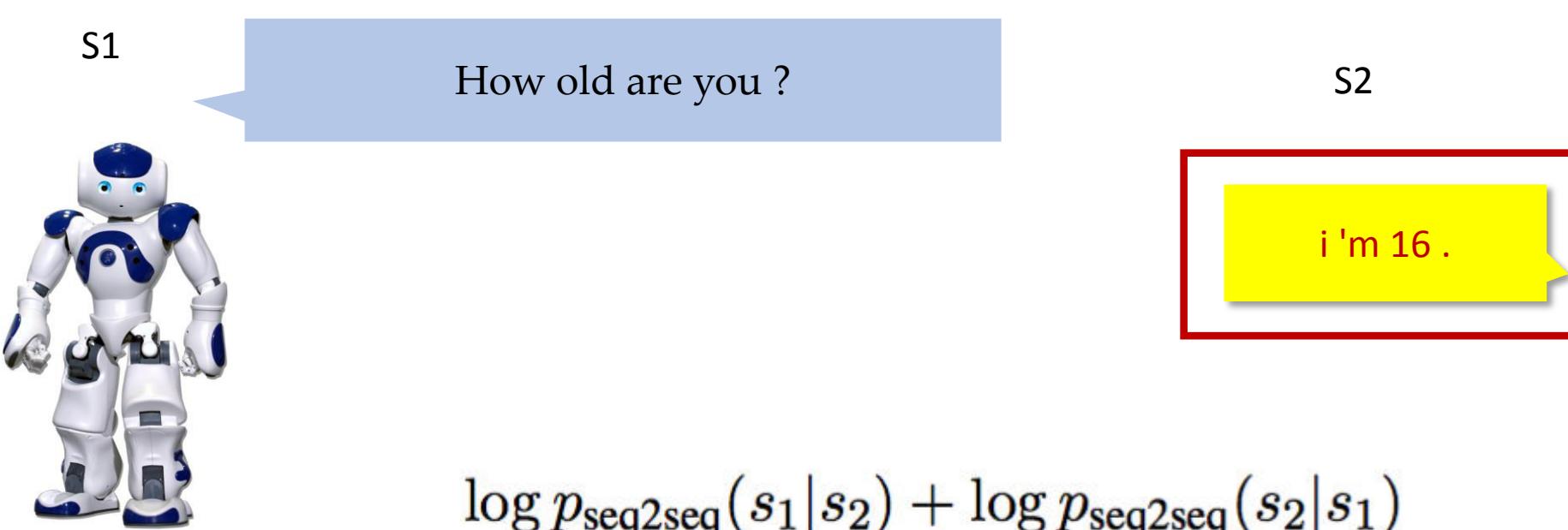
## 2. Information Flow



$$r_2 = -\log \text{Sigmoid}(\cos(s_1, s_2))$$

# Notations: Reward

## 3. Meaningfulness



# Notations: Reward

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

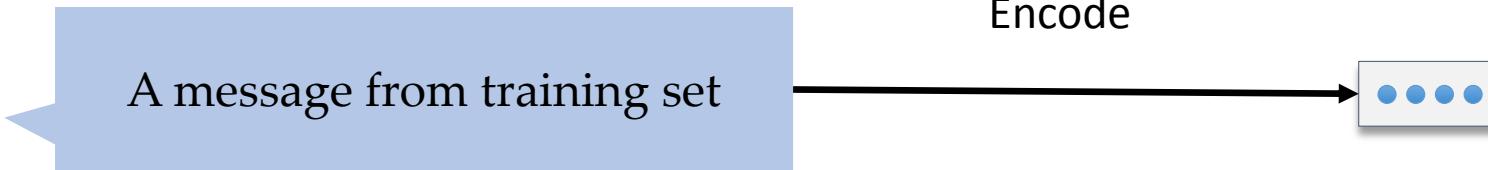
$$R = \lambda_1 R_1 + \lambda_2 R_2 + \lambda_3 R_3$$

# Simulation

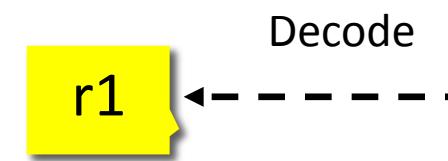
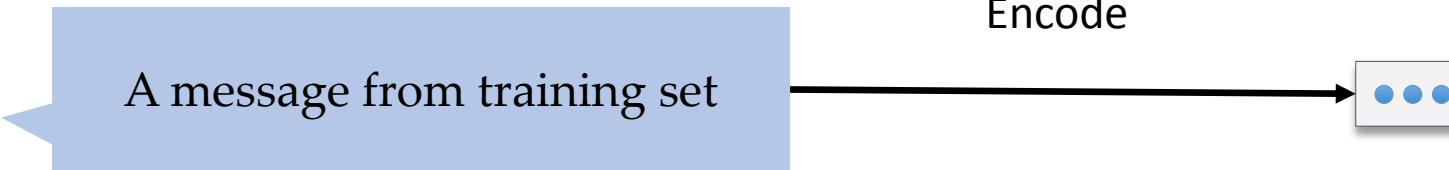


A message from training set

# Simulation

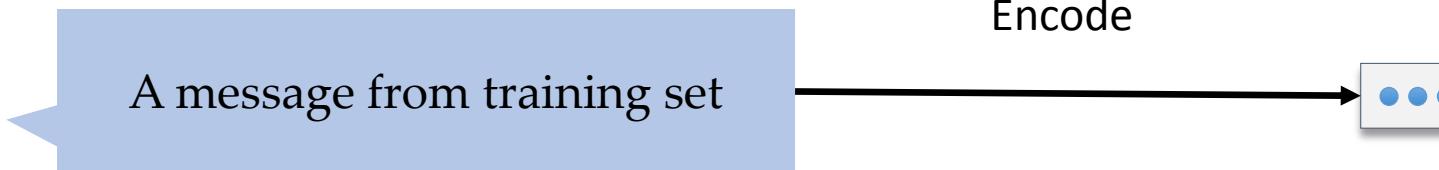


# Simulation



...

# Simulation



...

# Simulation



A message from training set

Encode



Encode



r1

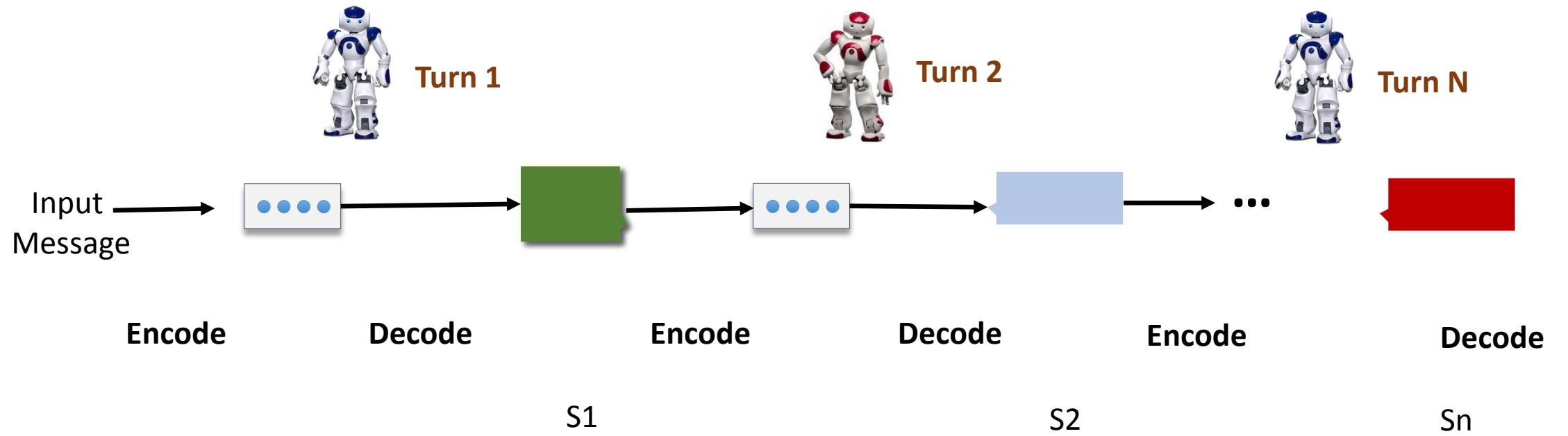
Decode

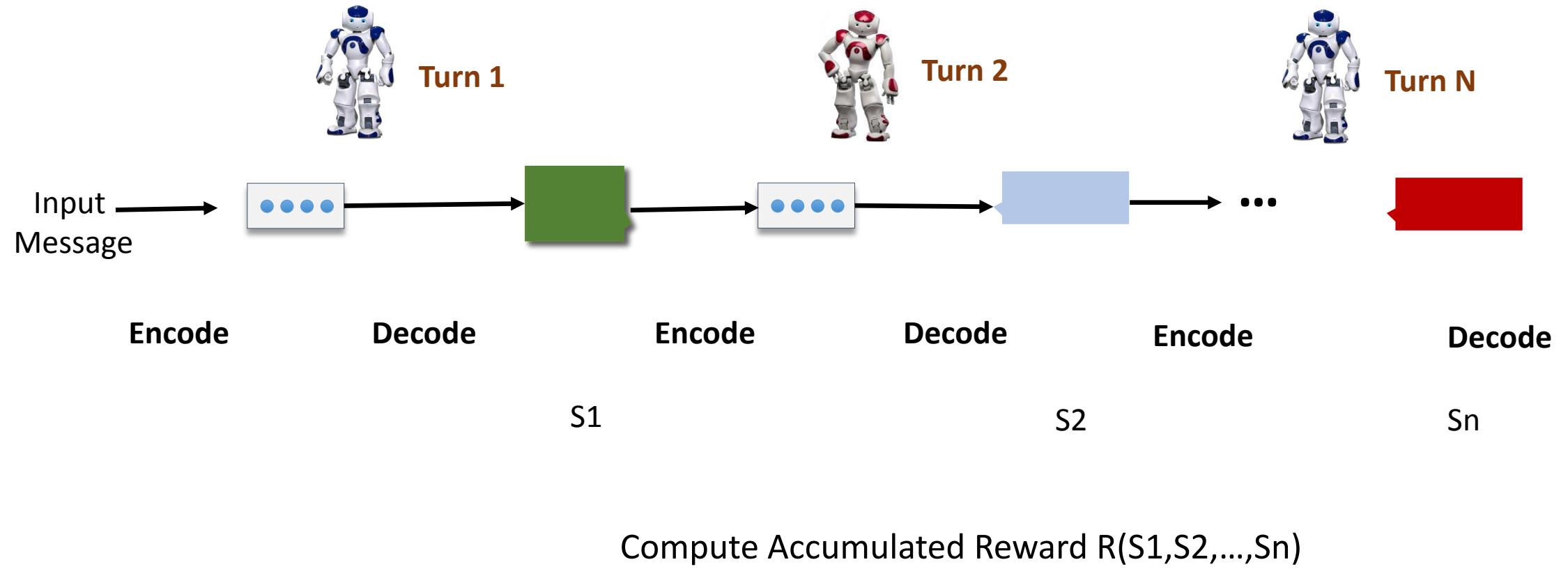


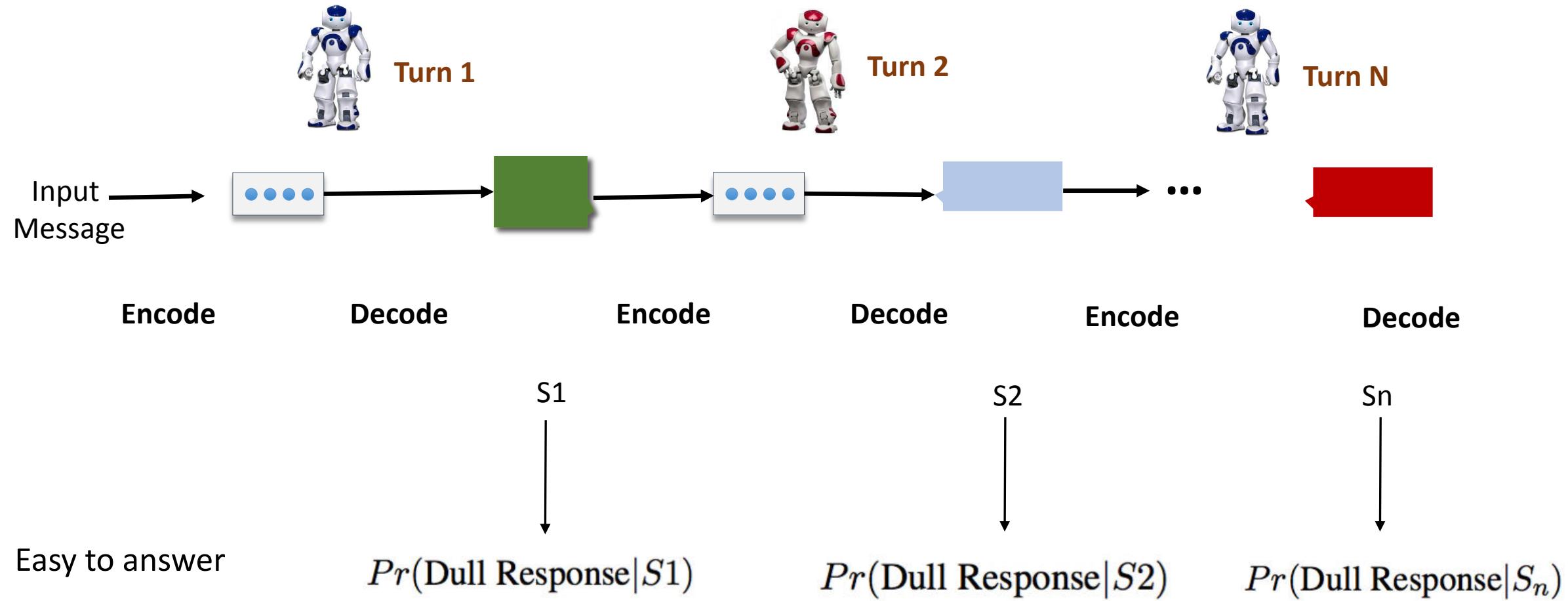
Decode

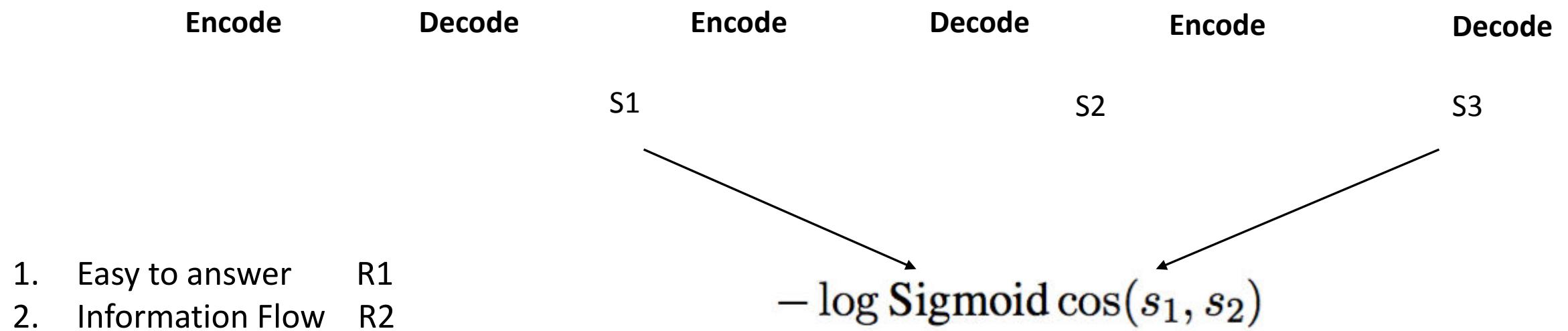
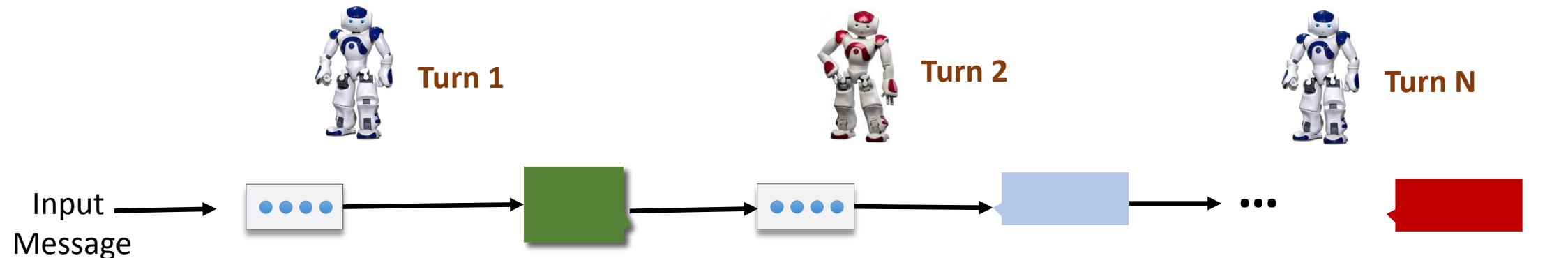
r2

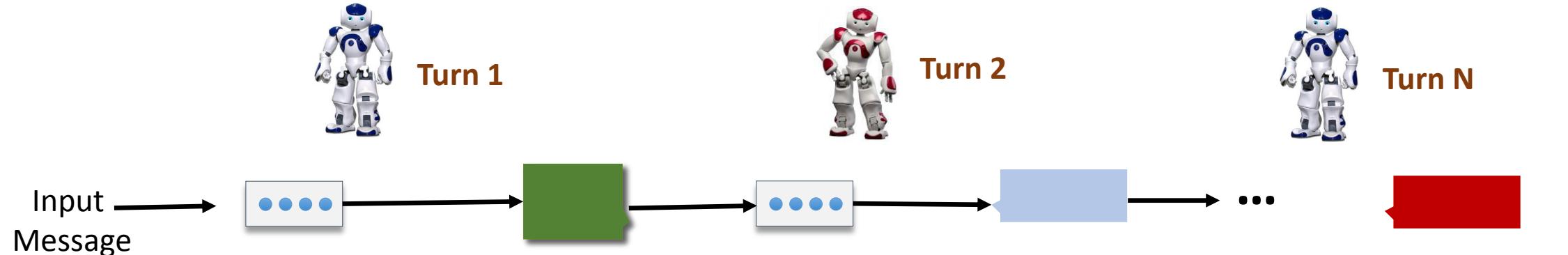
...









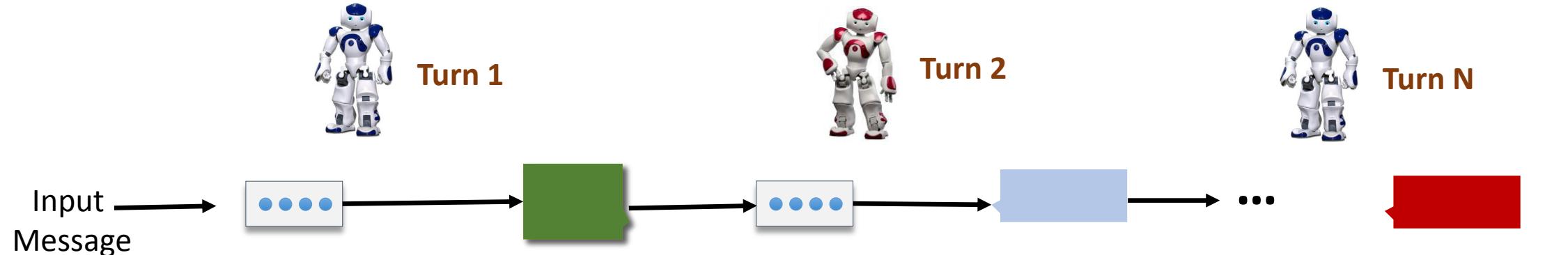


Encode      Decode      Encode      Decode      Encode      Decode

$s_1$                            $s_2$

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

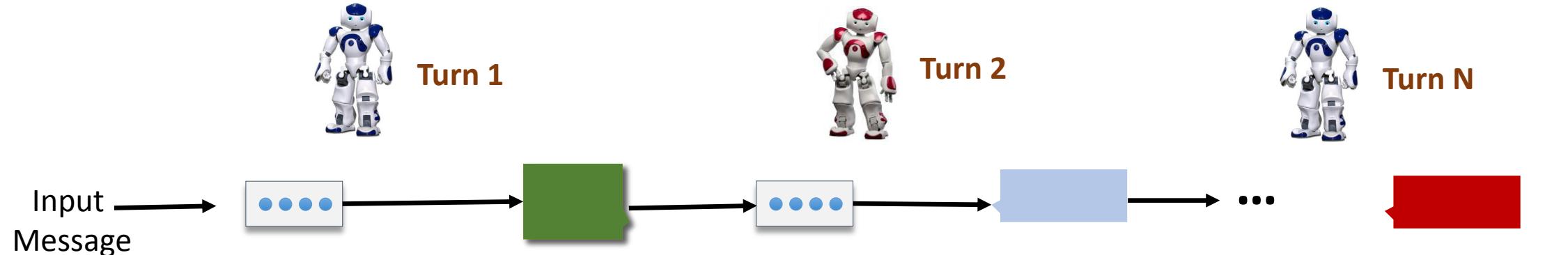
- |                     |    |
|---------------------|----|
| 1. Easy to answer   | R1 |
| 2. Information Flow | R2 |
| 3. Meaningfulness   | R3 |



**Encode**      **Decode**      **Encode**      **Decode**      **Encode**      **Decode**

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

$s_1 \rightarrow \log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1)$   
 $s_2 \rightarrow \log p_{\text{seq2seq}}(s_2|s_3) + \log p_{\text{seq2seq}}(s_3|s_2)$

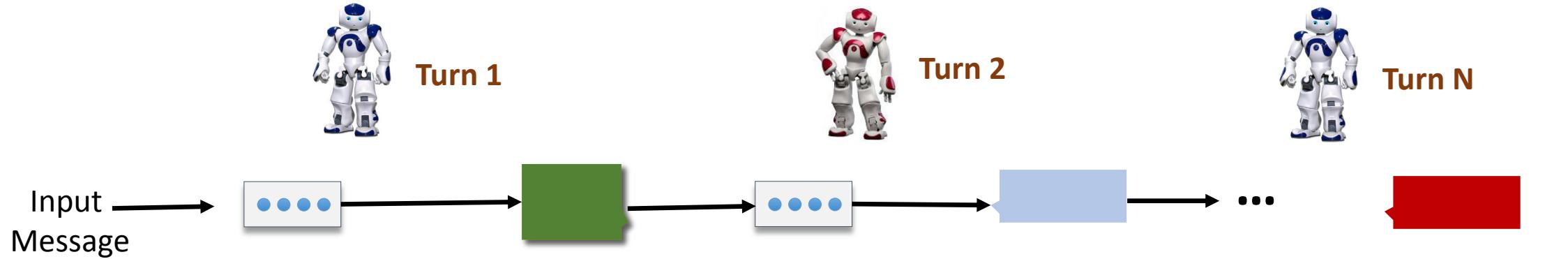


Encode      Decode      Encode      Decode      Encode      Decode

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

$s_1 \rightarrow \log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1)$

$s_2 \rightarrow \log p_{\text{seq2seq}}(s_2|s_3) + \log p_{\text{seq2seq}}(s_3|s_2)$

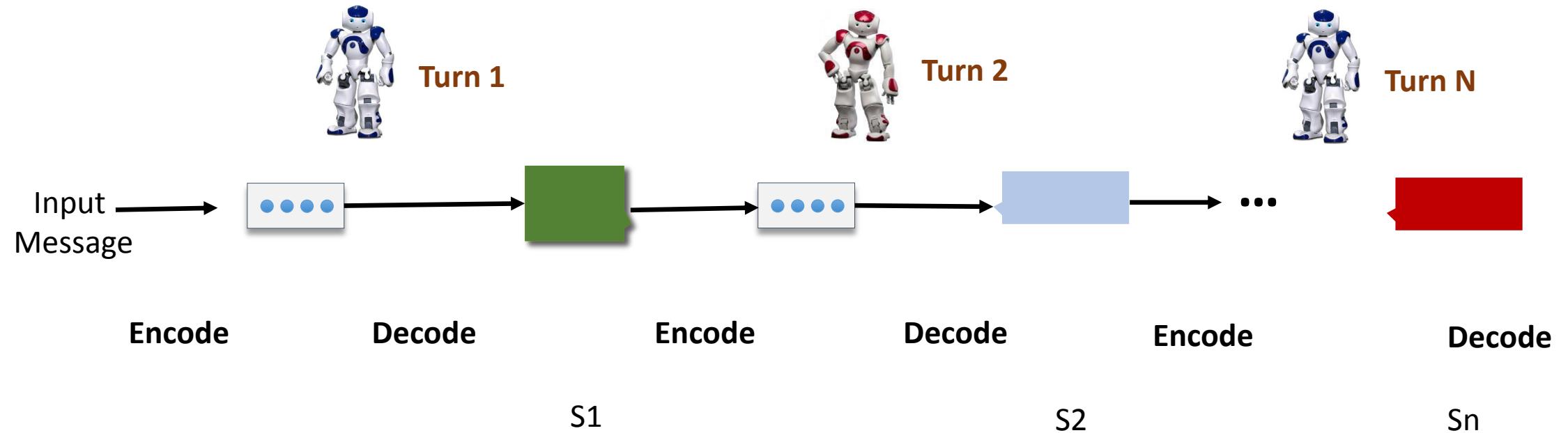


Encode      Decode      Encode      Decode      Encode      Decode

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

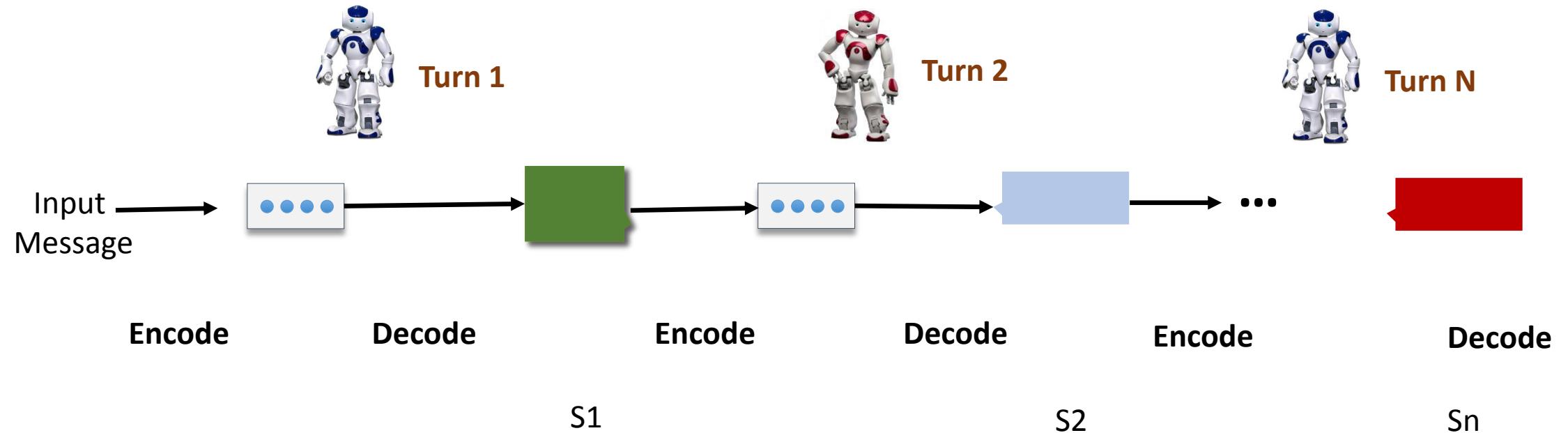
$$\begin{aligned}
 & \text{S1} \quad \log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1) \\
 & \text{S2} \quad \log p_{\text{seq2seq}}(s_2|s_3) + \log p_{\text{seq2seq}}(s_3|s_2)
 \end{aligned}$$

Compute Accumulated Reward R(S1,S2,...,Sn)



REINFORCE Algorithm (William, 1992)

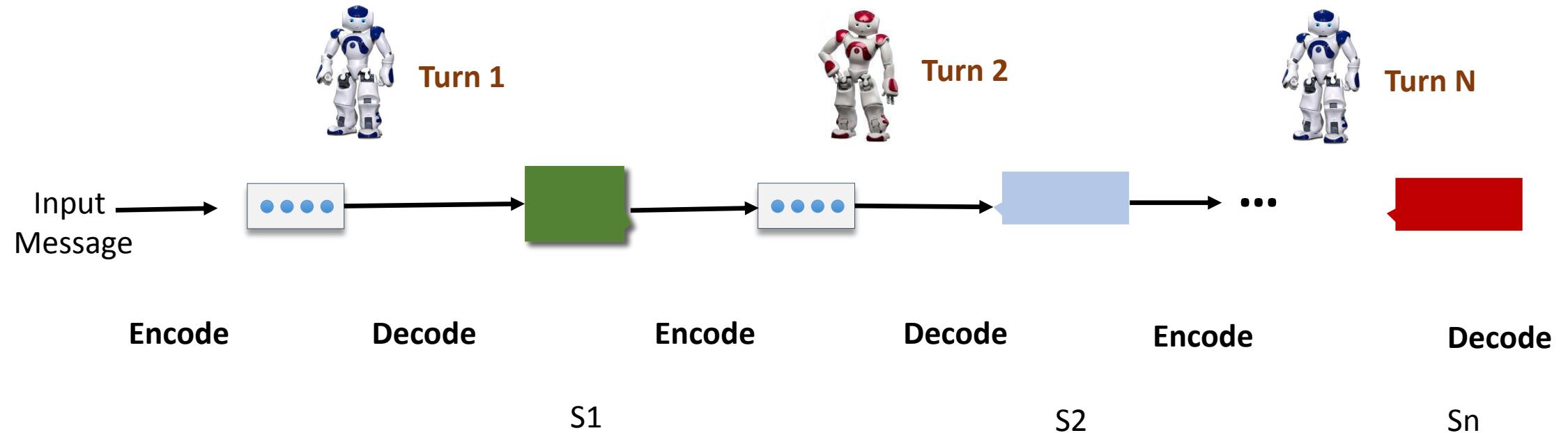
$$J(\theta) = \mathbb{E}[R(s_1, s_2, \dots, s_N)]$$



REINFORCE Algorithm (William, 1992)

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

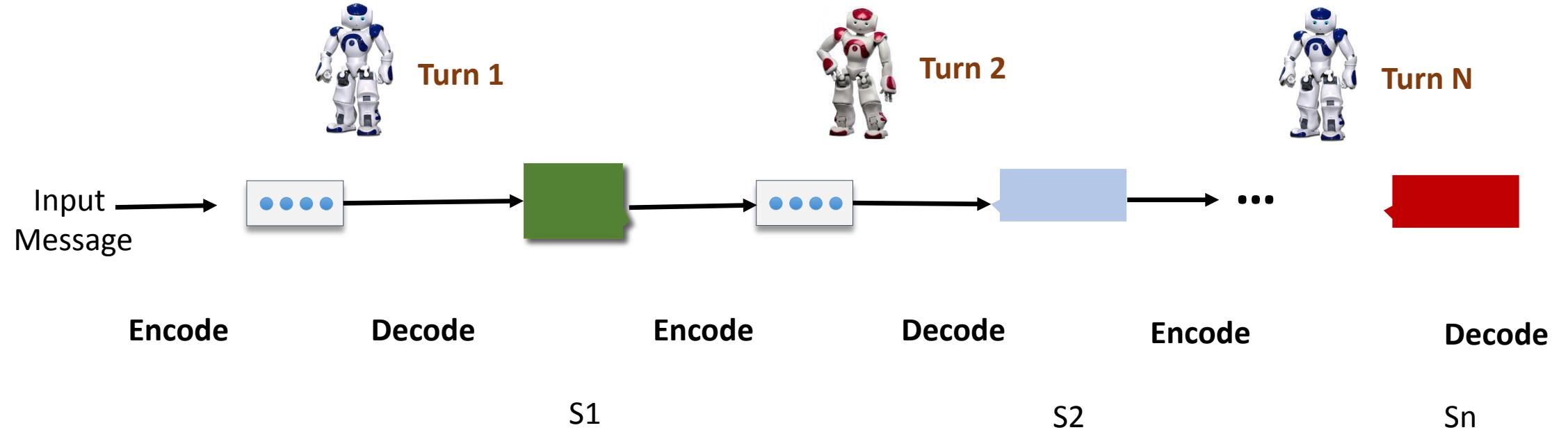
$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm (William, 1992)

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

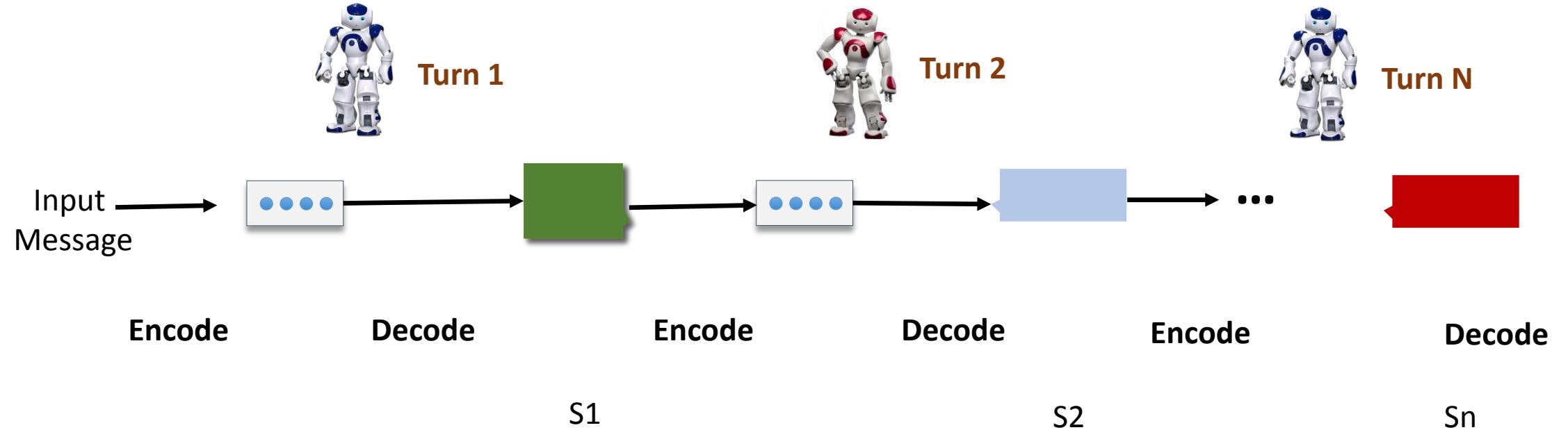
$$\nabla J(\theta) = \nabla \boxed{\log p(s_1, s_2, \dots, s_N)} R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm (William, 1992)

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

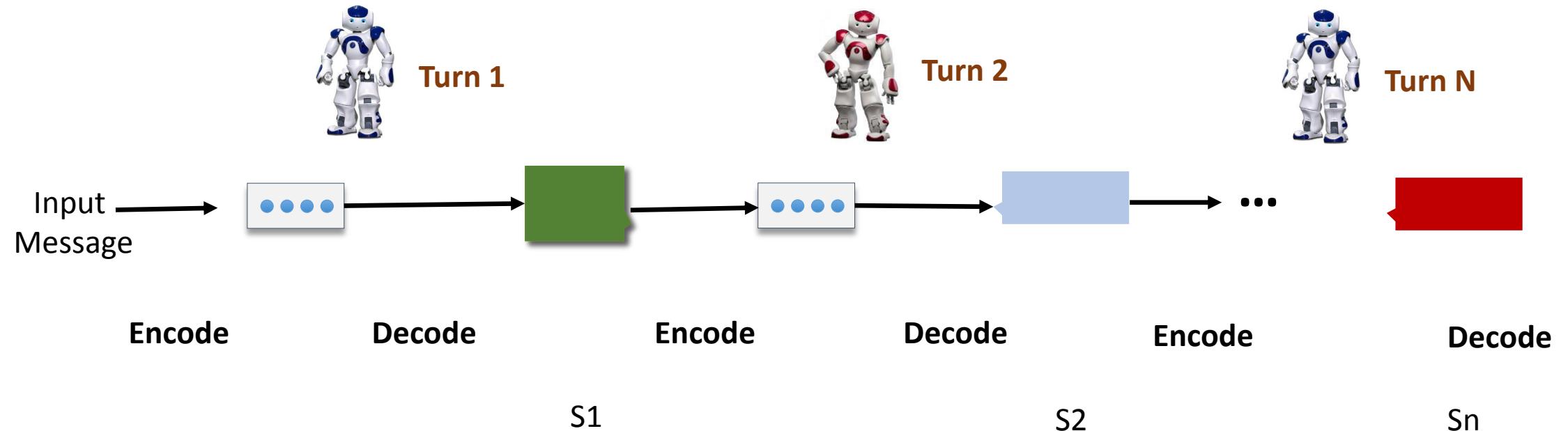
$$\nabla J(\theta) = \nabla \boxed{\log p(s_1, s_2, \dots, s_N)} R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm (William, 1992)

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

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

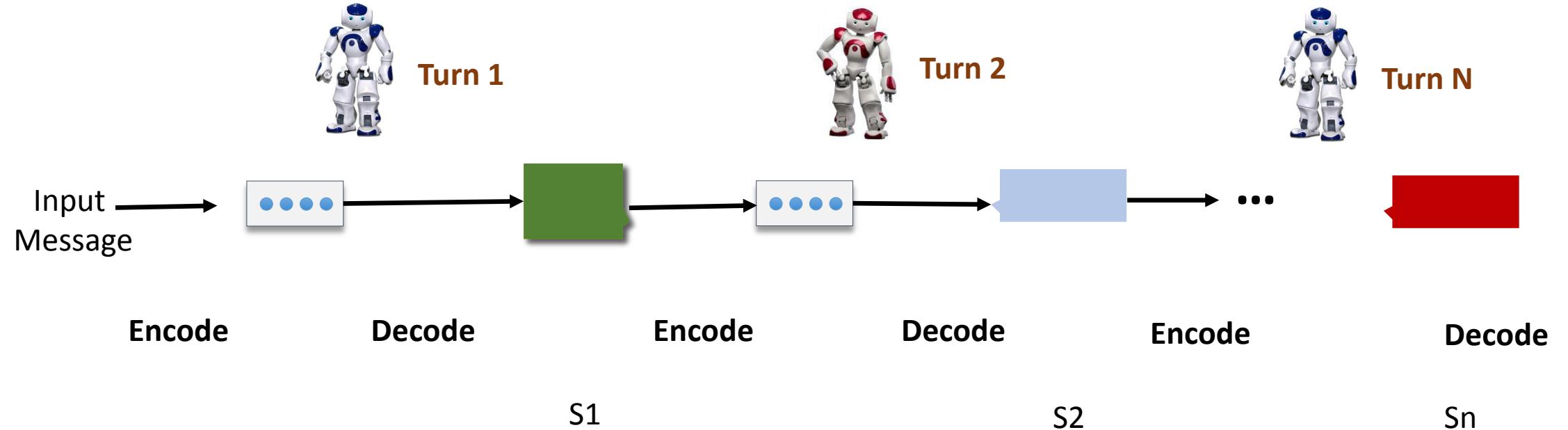


REINFORCE Algorithm (William, 1992)

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

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

$$\nabla J(\theta) = \nabla \log \prod_i p(s_i | s_{i-1}) R(s_1, s_2, \dots, s_N)$$



REINFORCE Algorithm (William, 1992)

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

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, \dots, s_N) R(s_1, s_2, \dots, s_N)$$

**What we want to learn**

$$\nabla J(\theta) = \nabla \log \prod_i p(s_i | s_{i-1}) R(s_1, s_2, \dots, s_N)$$

# Results

Input	Mutual Information	The proposed model
How old are you ?	I'm 16.	I'm 16. why are you asking

# Results

Input	Mutual Information	The proposed model
How old are you ?	I'm 16.	I'm 16. why are you asking ?

# Results

Input	Mutual Information	The proposed model
How old are you ?	I'm 16.	I'm 16. why are you asking ?
what is your full name ?	i have no idea	what 's yours ?
I don 't want to go home	Really ?	Why ?

# Results

Input	Mutual Information	The proposed model
How old are you ?	I'm 16.	I'm 16. 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



How old are you ?

# Simulation



How old are you ?

i 'm 16, why are  
you asking ?



# Simulation



How old are you ?

I thought you were 12 .

i 'm 16, why are  
you asking ?



# Simulation



How old are you ?

I thought you were 12 .

i 'm 16, why are  
you asking ?

What made you think so ?



# Simulation



How old are you ?

I thought you were 12 .

I don't know what you are talking about .

i 'm 16. why are you asking ?

What made you think so ?



# Simulation



How old are you ?

i 'm 16, 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



How old are you ?

I thought you were 12 .

I don't know what you are talking about .

I don't know what you are talking about .

i 'm 16, why are you asking ?

What made you think so ?

You don't know what you are saying.



# Simulation

Survive 4 turns !!



How old are you ?

I thought you were 12 .

I don't know what you are talking about .

I don't know what you are talking about .

i 'm 16, why are you asking ?

What made you think so ?

You don't know what you are saying.



# Outline

# Outline

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

# Reward for Good Dialogue

# Reward for Good Dialogue



How to evaluate open  
domain dialogue generation

# Reward for Good Dialogue

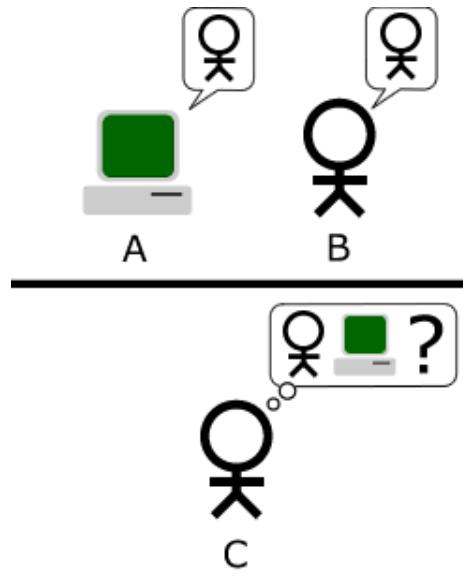


How to evaluate open  
domain dialogue generation

Bleu ? Ppl ?

# Reward for Good Dialogue

Turing Test

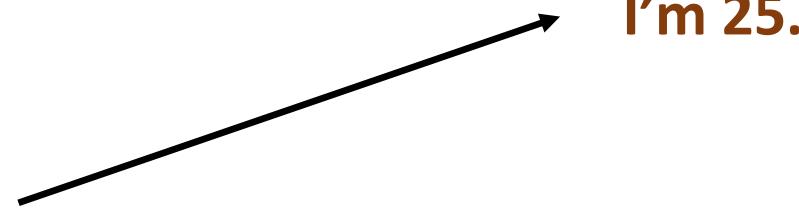


# Reward for Good Dialogue

## Turing Test

Input

How old are you ?



I'm 25.

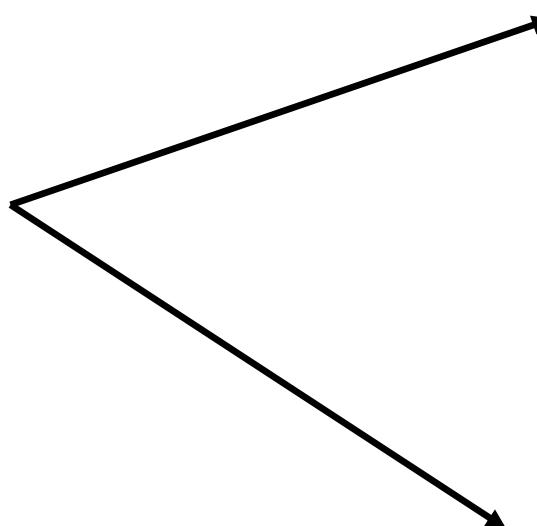


# Reward for Good Dialogue

## Turing Test

Input

How old are you ?



I'm 25.



I don't know what you are  
talking about



# Reward for Good Dialogue

## Turing Test

How old are you ?



A human evaluator/ judge

I'm 25.



I don't know what you are  
talking about



# Reward for Good Dialogue

## Turing Test

How old are you ?



A human evaluator/ judge

I'm 25.



I don't know what you are  
talking about



# Reward for Good Dialogue

## Turing Test

How old are you ?



To expensive ?

I'm 25.



I don't know what you are  
talking about



# Reward for Good Dialogue

How old are you ?



I'm 25.

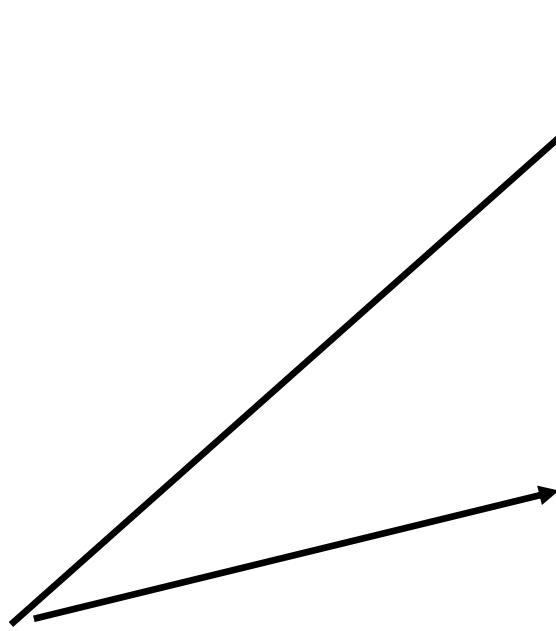


I don't know what you are  
talking about



# Reward for Good Dialogue

How old are you ?



P= 90% human generated

I'm 25.



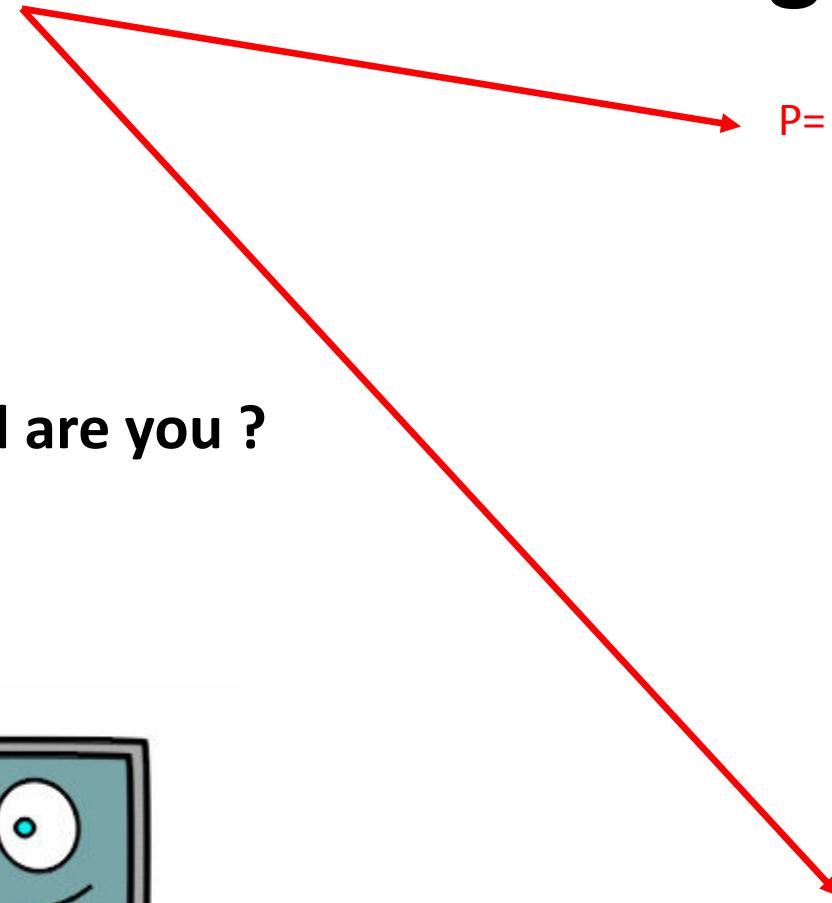
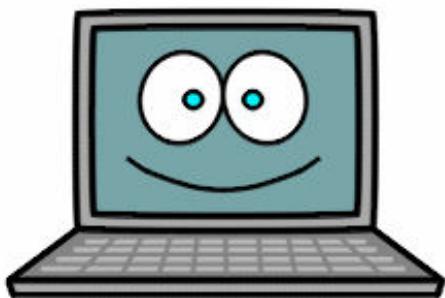
I don't know what you are  
talking about

P= 10% human generated



# Reward for Good Dialogue

How old are you ?



P= 90% human generated

I'm 25.

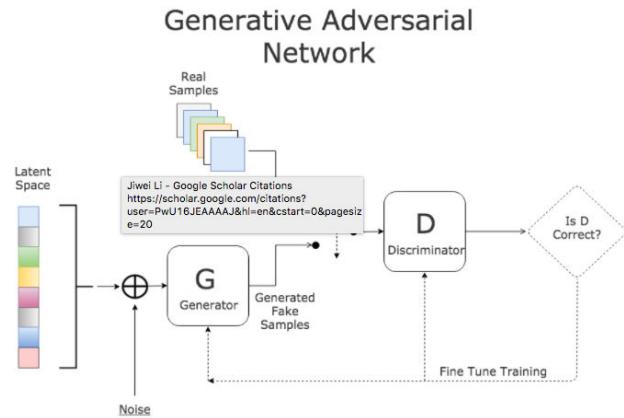


I don't know what you are  
talking about

P= 10% human generated



# Adversarial Learning in Image Generation (Goodfellow et al., 2014)

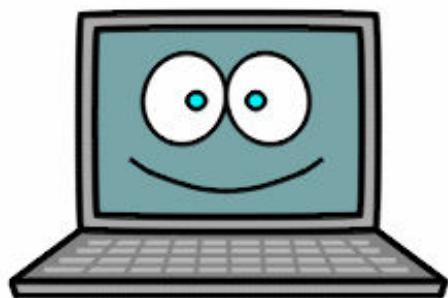


P= 90% human generated

I'm 25.



Discriminator

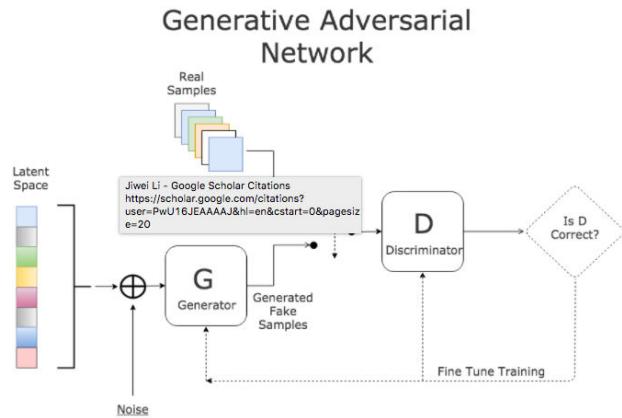


I don't know what you are talking about

P= 10% human generated



# Adversarial Learning in Image Generation (Goodfellow et al., 2014)



Discriminator



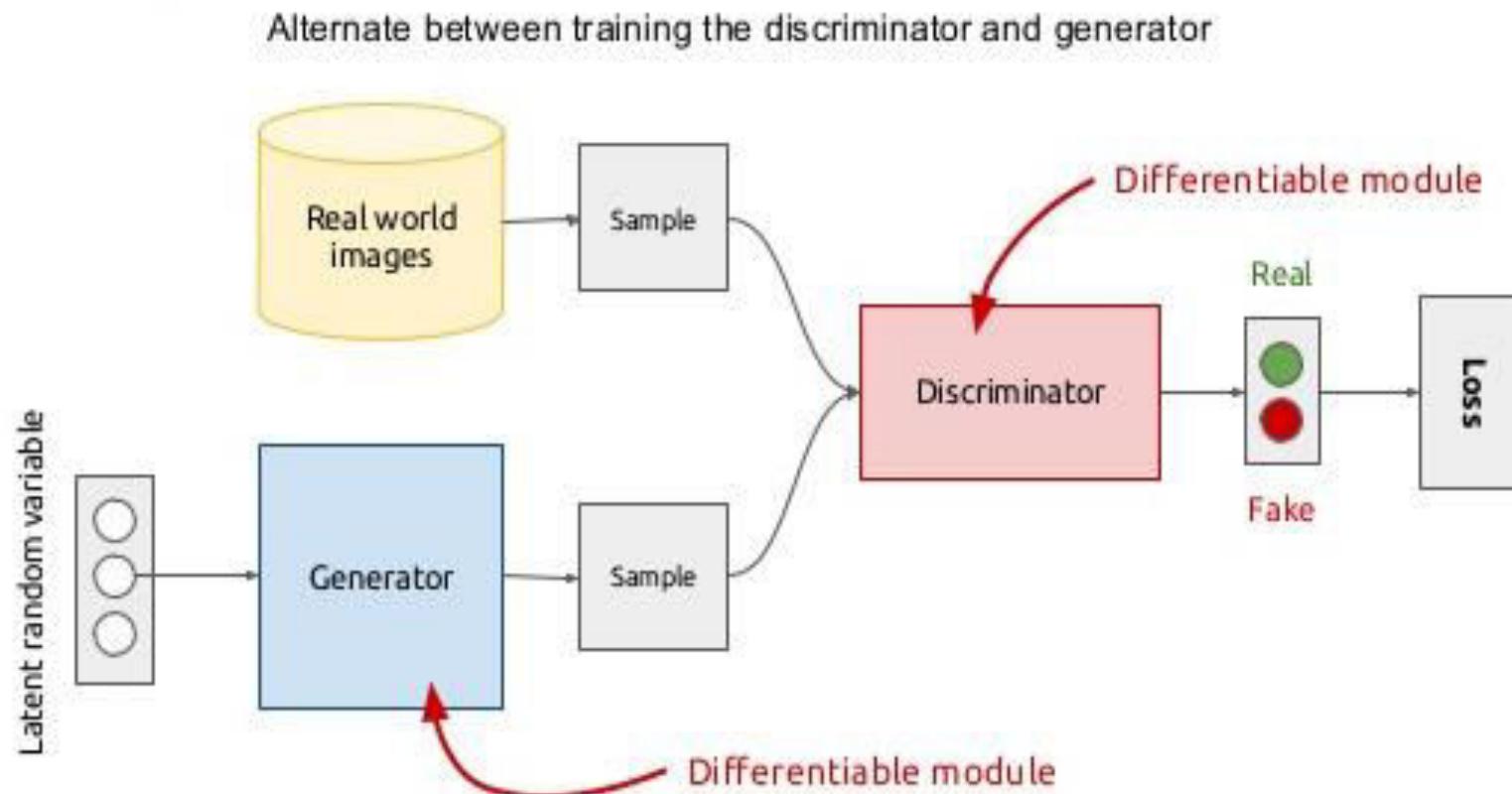
P= 90% human generated



P= 10% human generated



# Adversarial Learning in Image Generation (Goodfellow et al., 2014)



# Model Break Down

Generative Model (G)

Discriminative Model (D)

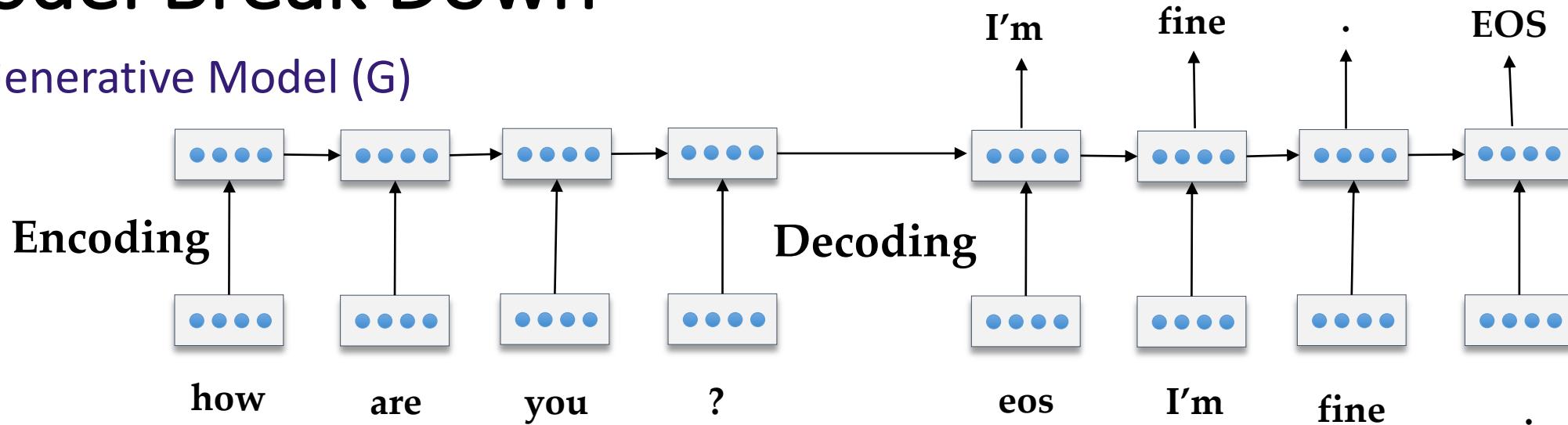
# Model Break Down

Generative Model (G)

Discriminative Model (D)

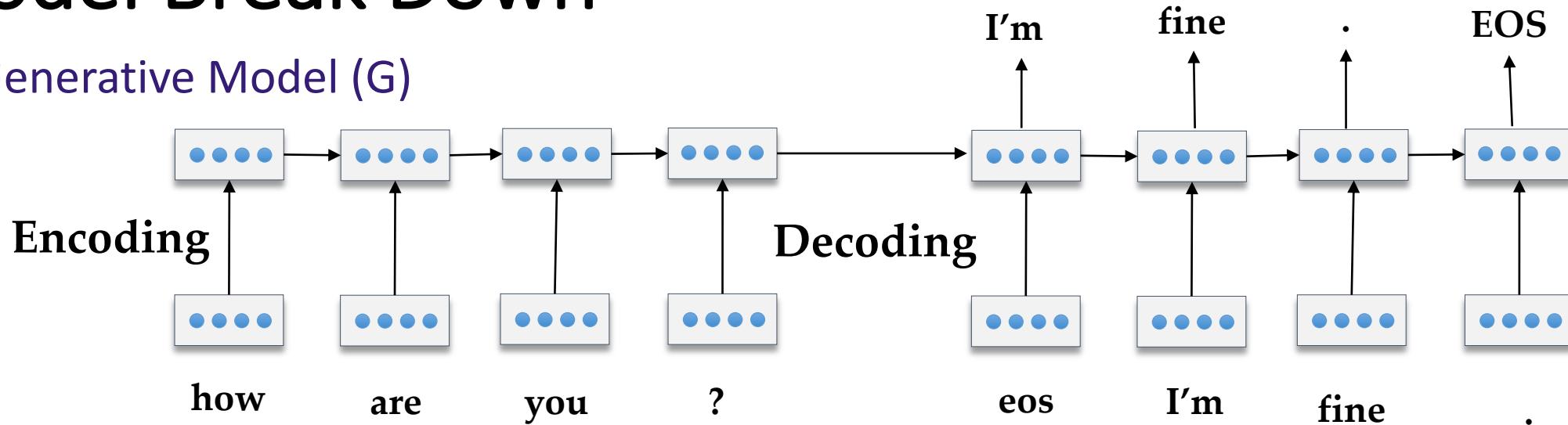
# Model Break Down

Generative Model (G)



# Model Break Down

Generative Model (G)

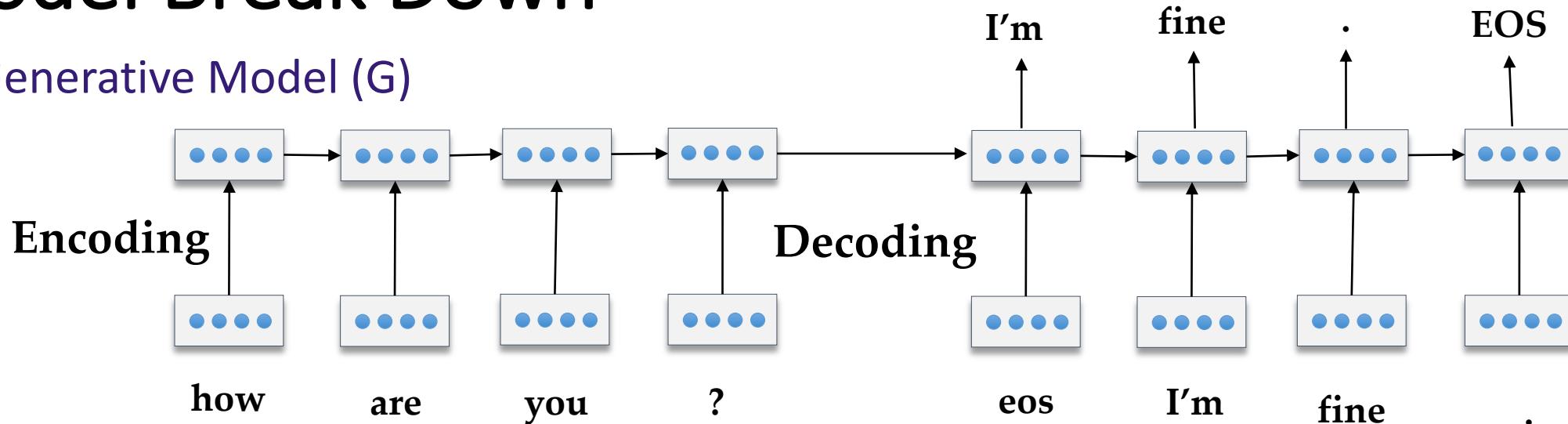


Discriminative Model (D)

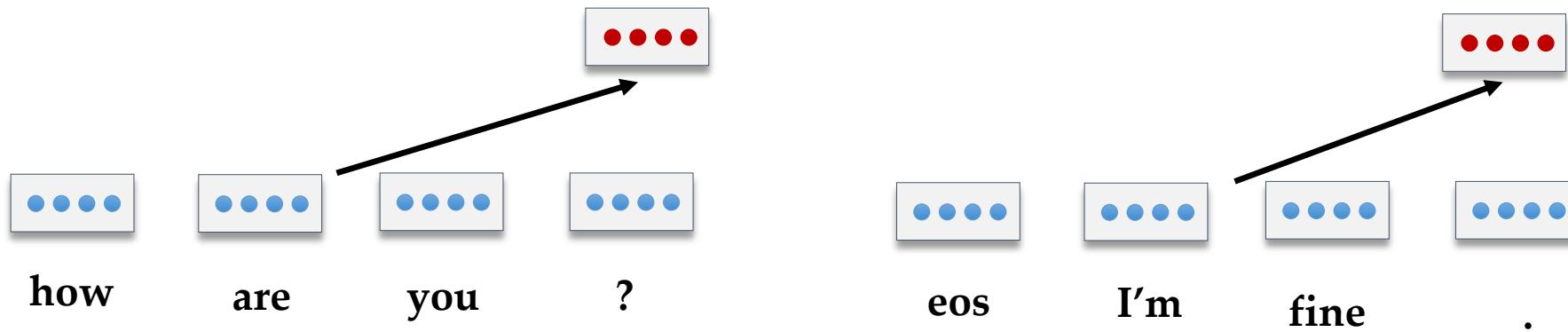


# Model Break Down

Generative Model (G)

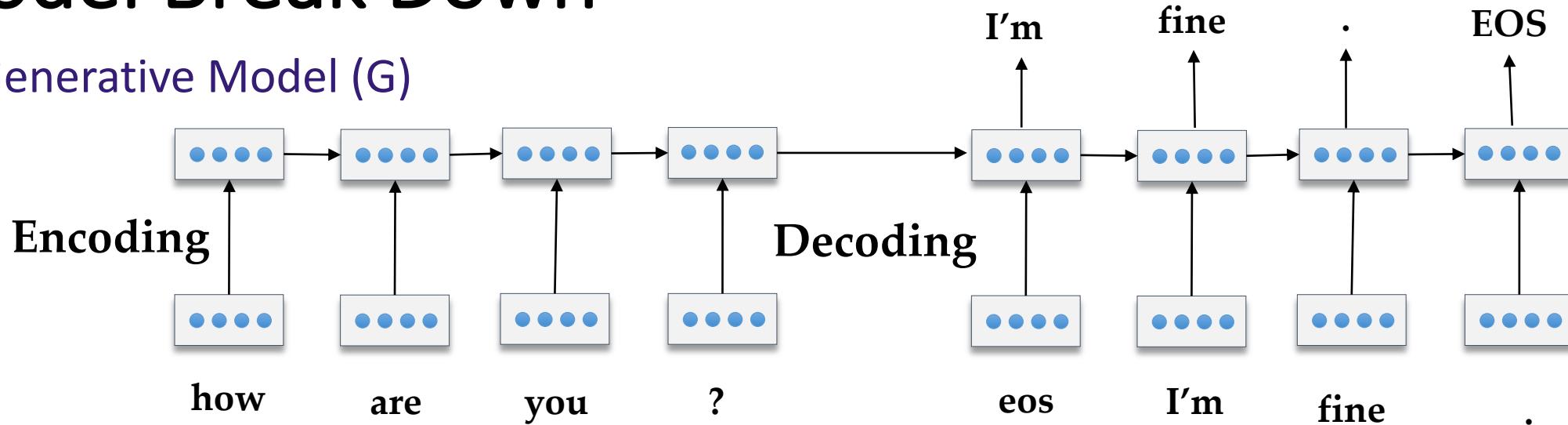


Discriminative Model (D)



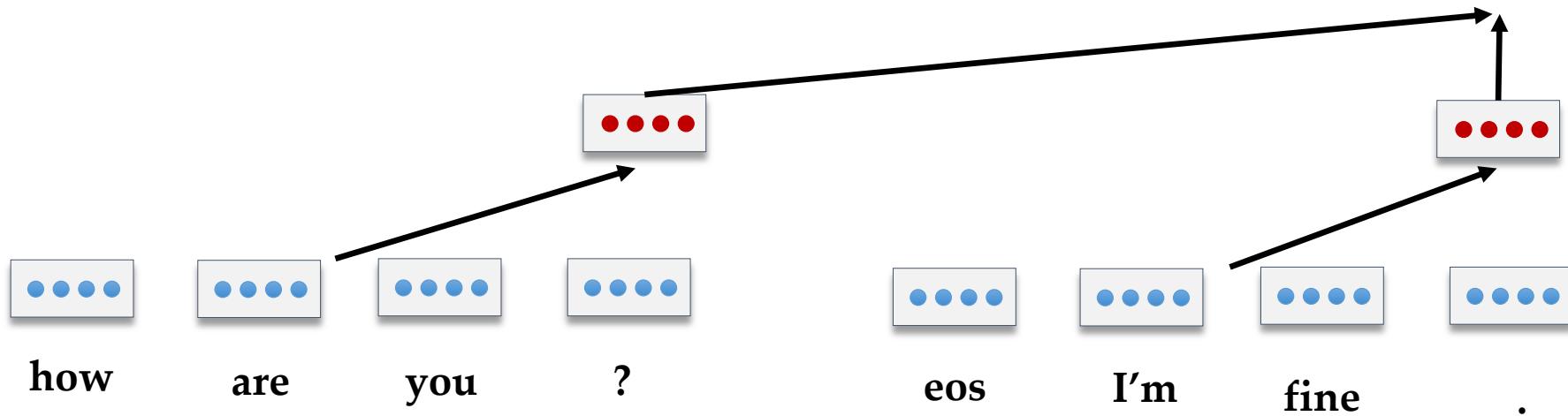
# Model Break Down

Generative Model (G)



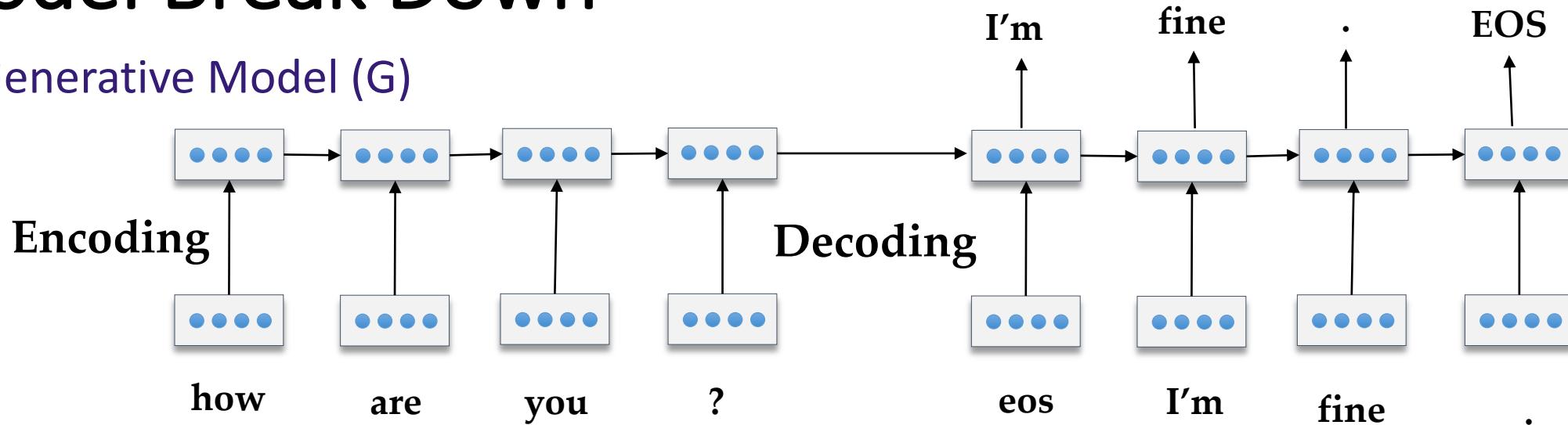
Discriminative Model (D)

P= 90% human generated

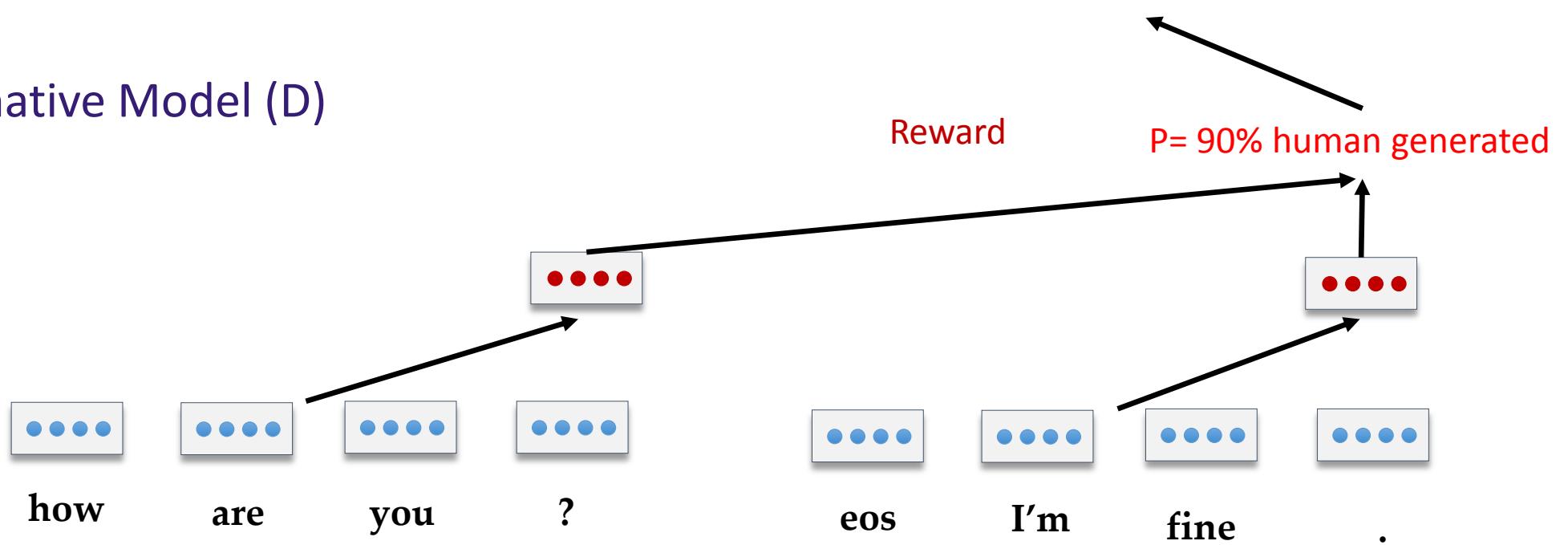


# Model Break Down

Generative Model (G)

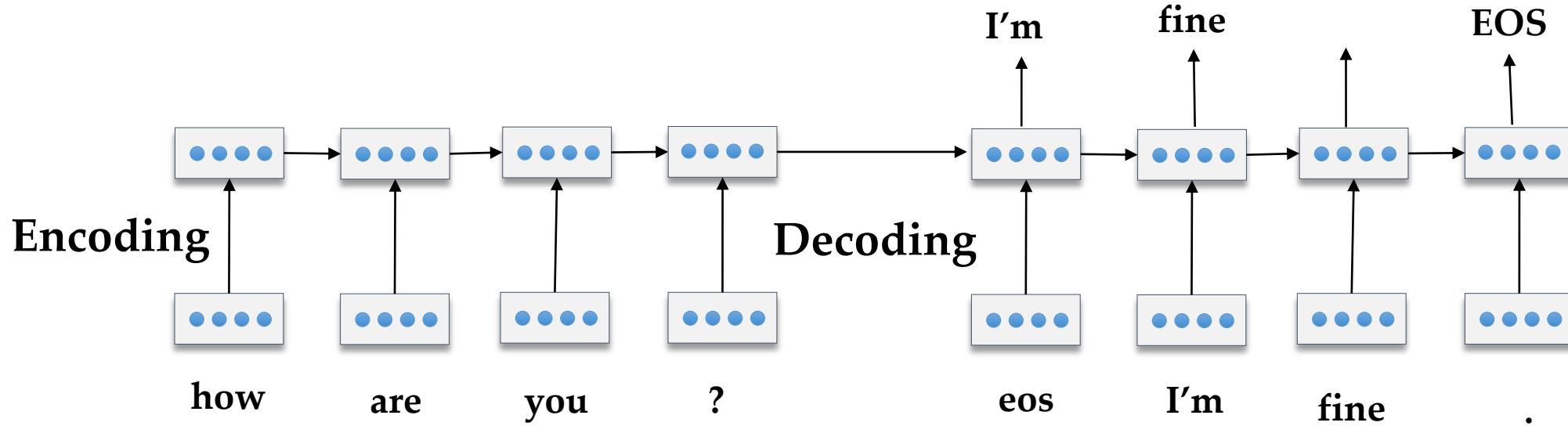


Discriminative Model (D)



# Policy Gradient

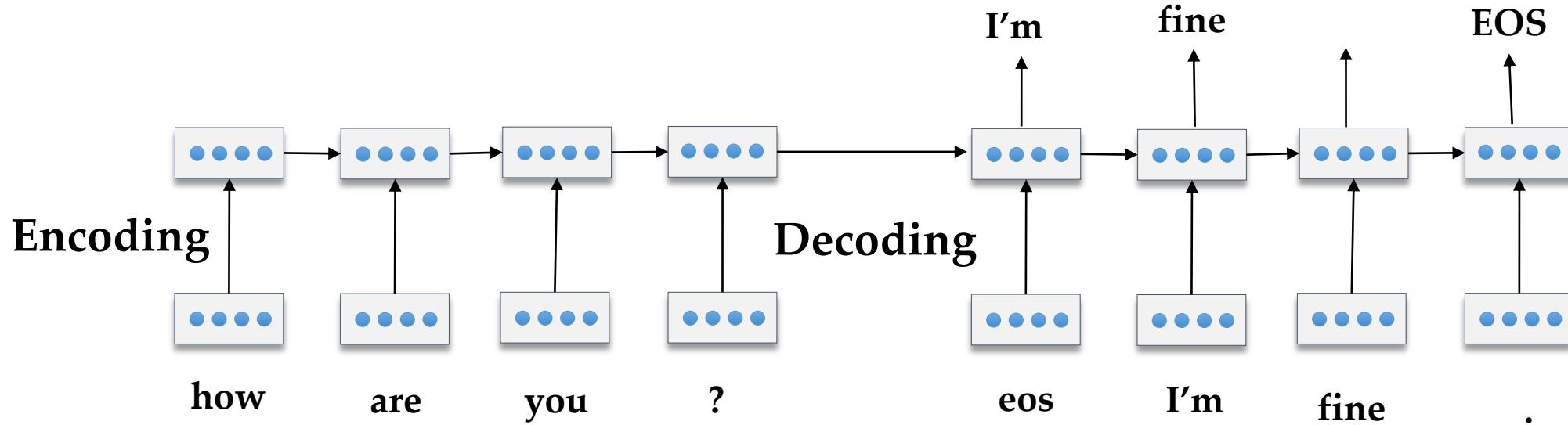
Generative Model (G)



REINFORCE Algorithm (William, 1992)

# Policy Gradient

Generative Model (G)

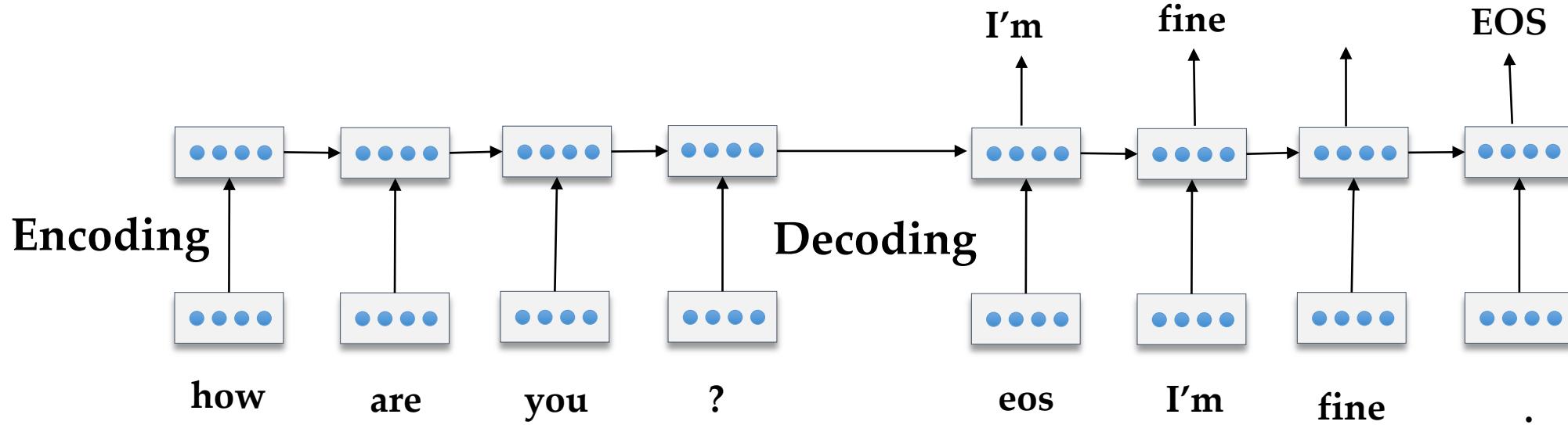


REINFORCE Algorithm (William, 1992)

$$J = E[R(y)]$$

# Policy Gradient

Generative Model (G)

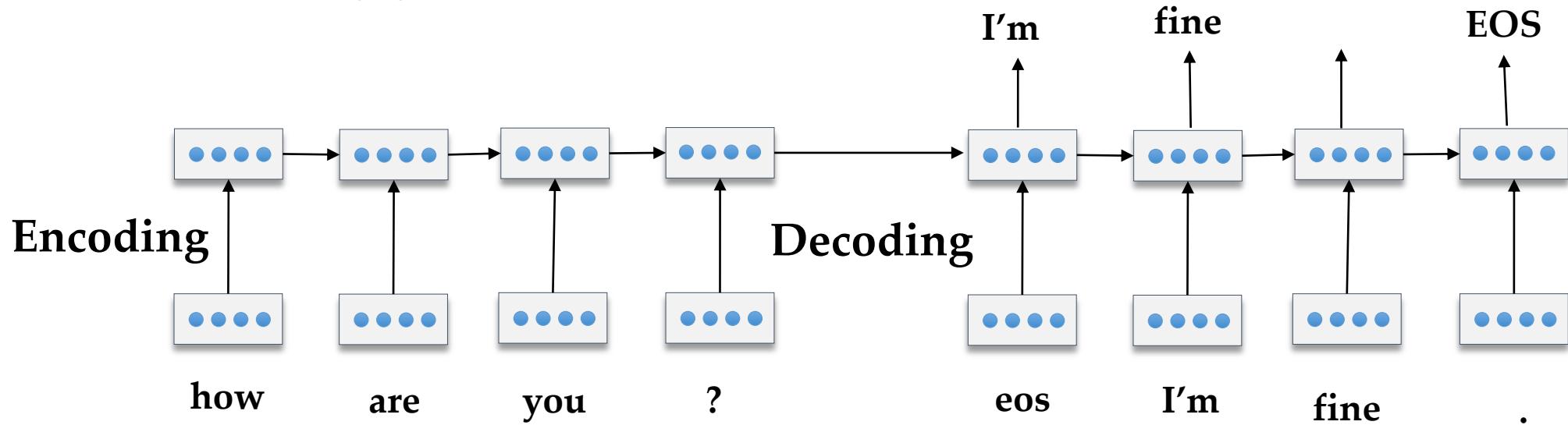


REINFORCE Algorithm (William, 1992)

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

# Policy Gradient

Generative Model (G)

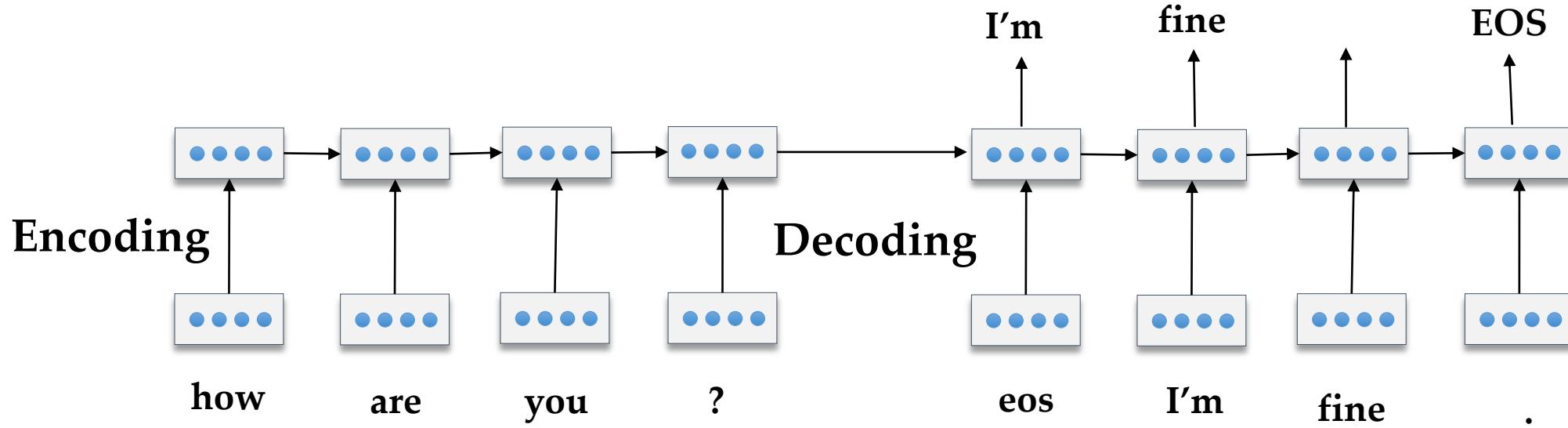


REINFORCE Algorithm (William, 1992)

$$\begin{aligned} J &= E[R(y)] \\ \nabla J &= R(y) \nabla \log p(y) \\ &= R(y) \sum_i \log p(y_i | x, y_{1, \dots, i-1}) \end{aligned}$$

# Policy Gradient

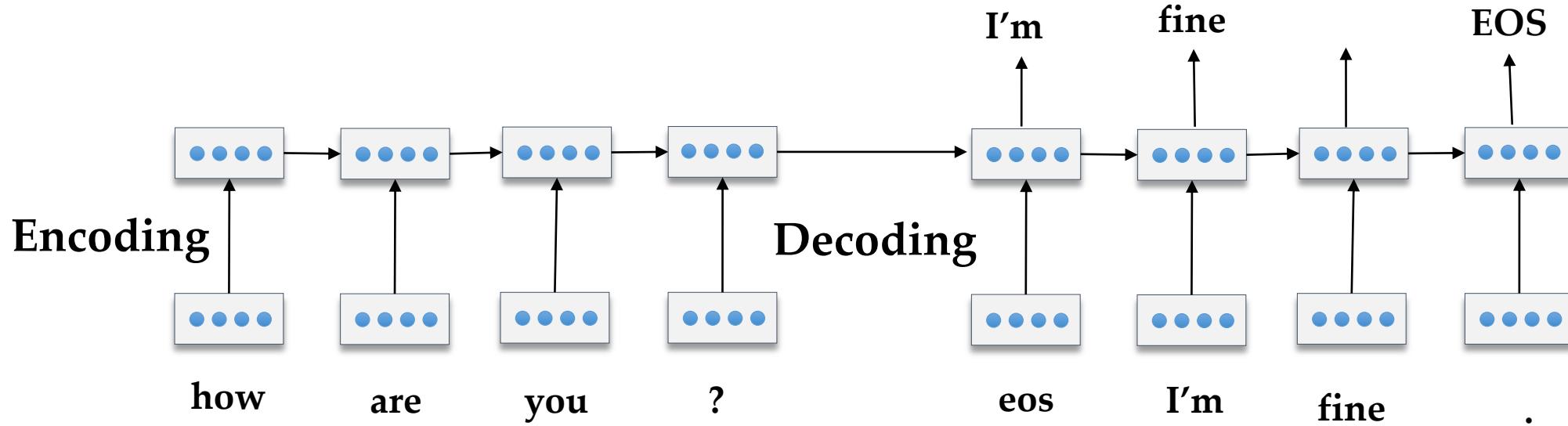
Generative Model (G)



$$\begin{aligned}\nabla J &= R(y) \nabla \log p(y) \\ &= R(y) \sum_i \log p(y_i | x, y_{1, \dots, i-1})\end{aligned}$$

# Policy Gradient

Generative Model (G)



$$\begin{aligned}\nabla J &= R(y) \nabla \log p(y) \\ &= R(y) \sum_i \log p(y_i | x, y_{1, \dots, i-1})\end{aligned}$$

Same reward to all tokens tokens

# Policy Gradient

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

# Policy Gradient

Input : *What's your name*

human : *I am John*

machine : *I don't know*

- REINFORCE assigns the same negative reward to all tokens [*I, don't, know*]

# Policy Gradient

Input : *What's your name*

human : *I am John*

machine : *I don't know*

- REINFORCE assigns the same negative reward to all tokens [*I, don't, know*]



Negative reward for “ I ”

# Policy Gradient

Input : *What's your name*

human : *I am John*

machine : *I don't know*

- REINFORCE assigns the same negative reward to all tokens [*I, don't, know*]

Reward for Every Generation Step (REGS)

# Policy Gradient

## Monte Carlo Search

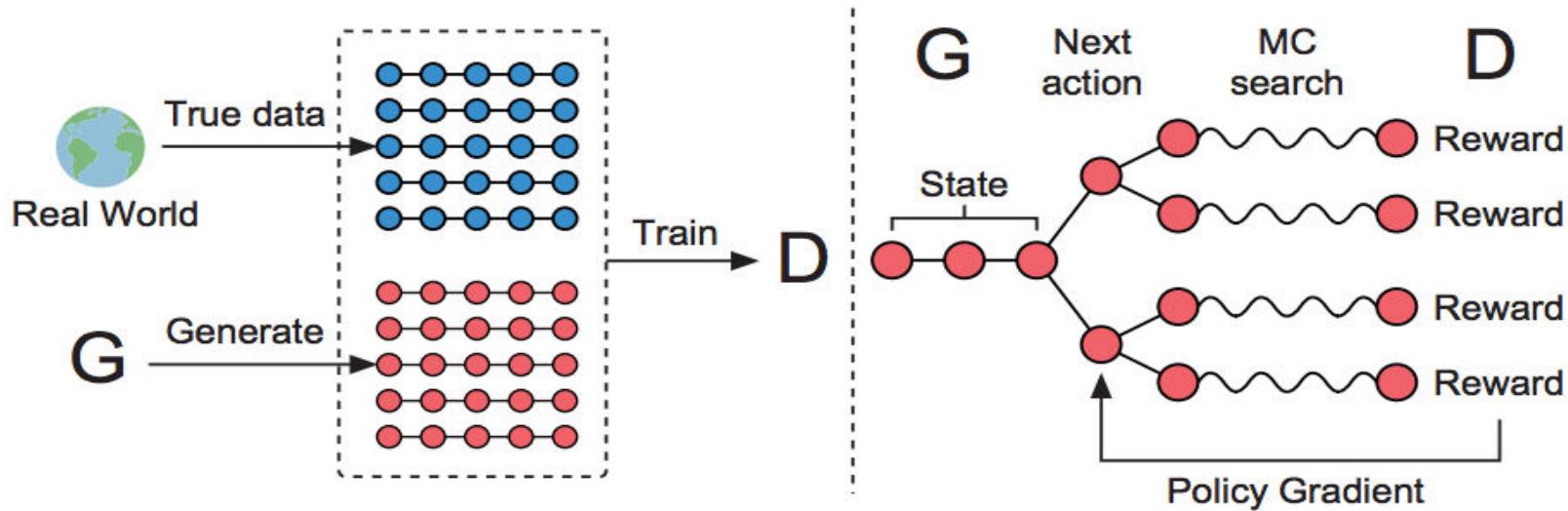
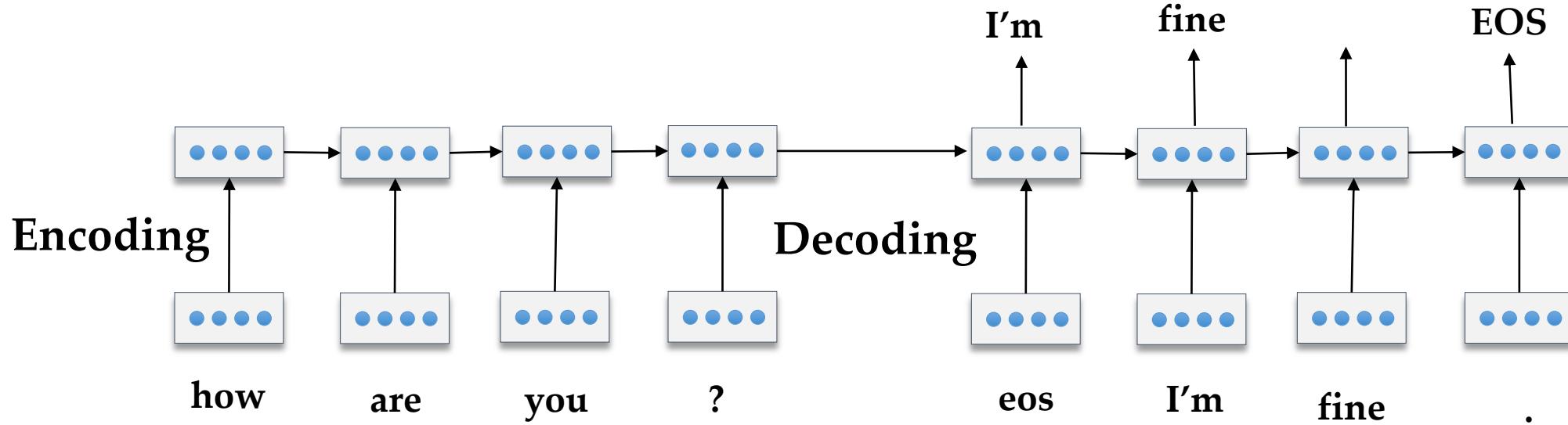


Figure from Yu et al (2016)

# Policy Gradient

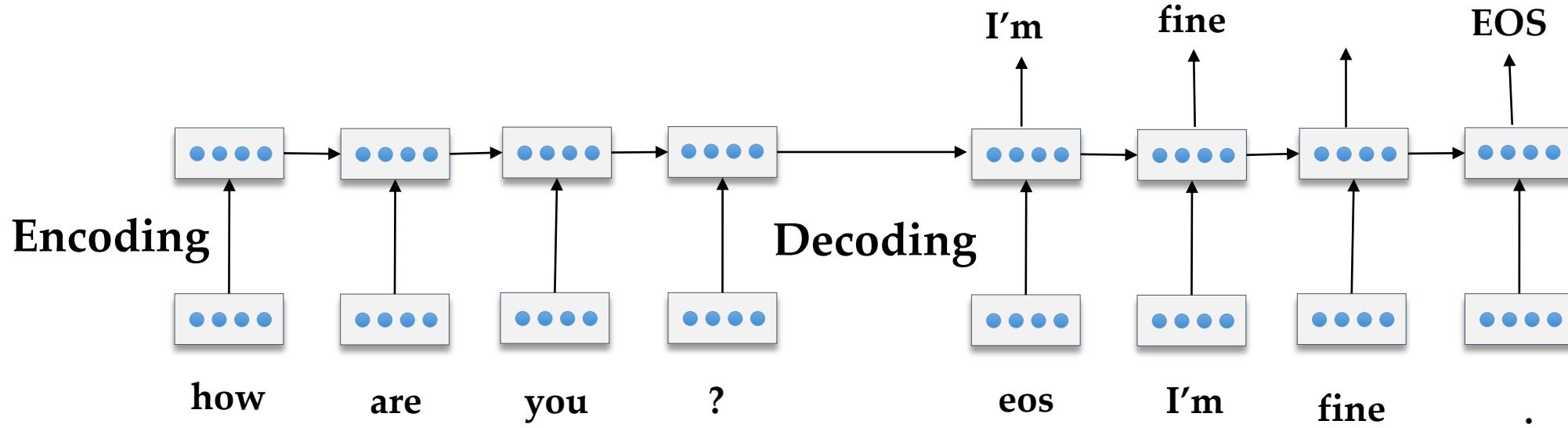
Generative Model (G)



$$\sum_i R(y_i) \log p(y_i|x, y_{1,\dots,i-1})$$

# Policy Gradient

Generative Model (G)

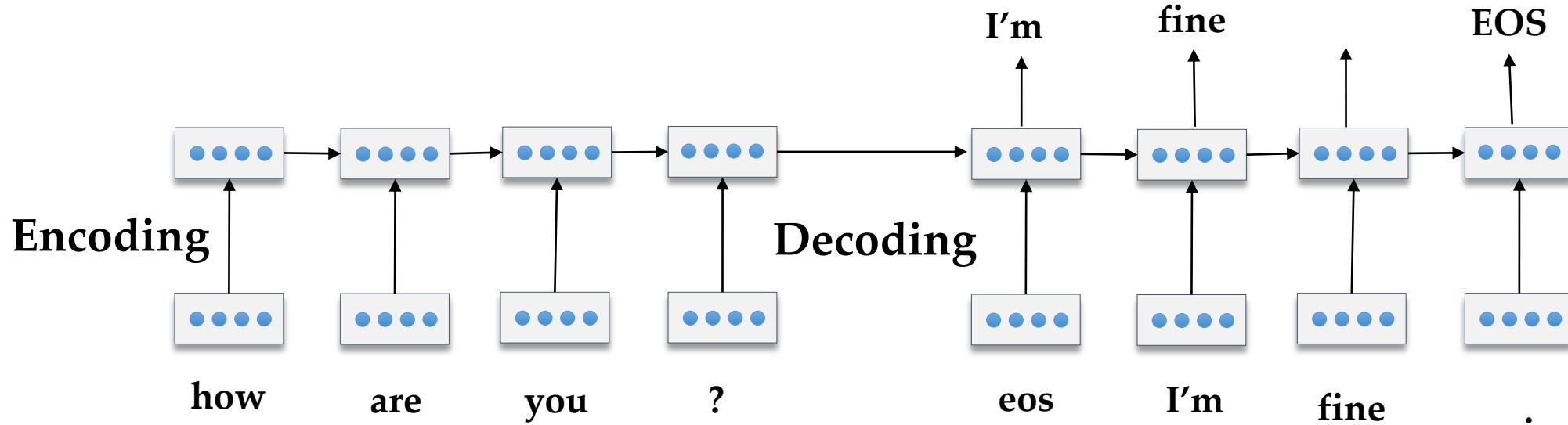


$$\sum_i R(y_i) \log p(y_i|x, y_{1,\dots,i-1}) - R(y) \sum_i \log p(y_i|x, y_{1,\dots,i-1})$$

Difference

# Policy Gradient

Generative Model (G)



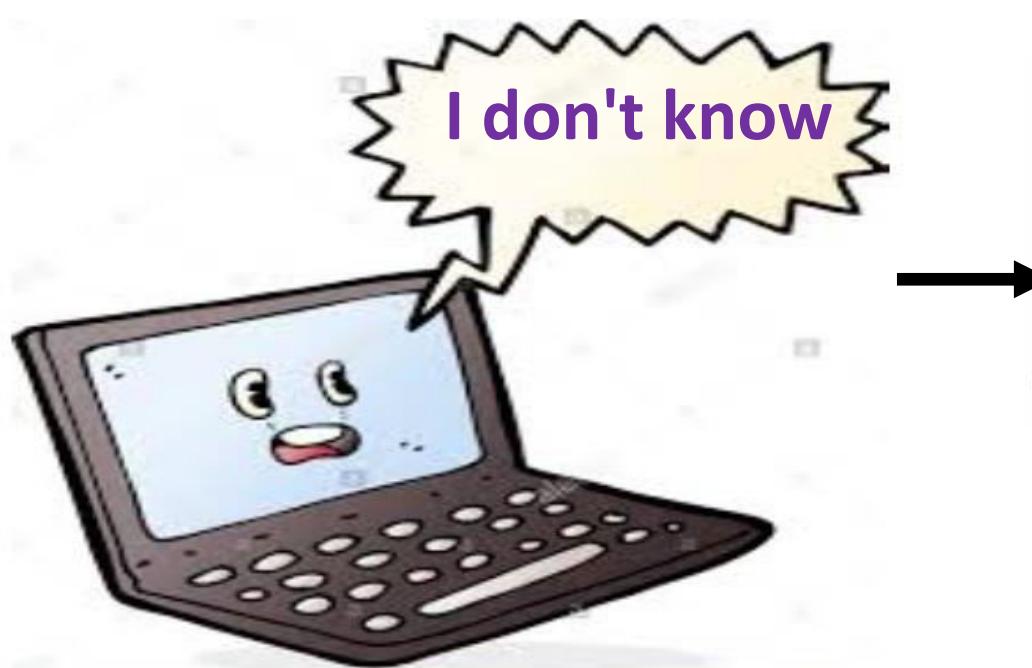
$$\sum_i R(y_i) \log p(y_i|x, y_{1,\dots,i-1})$$

$$\sum_i [R(y_i) - b(y_i)] \log p(y_i|x, y_{1,\dots,i-1}) \quad \text{Actor-critic RL}$$

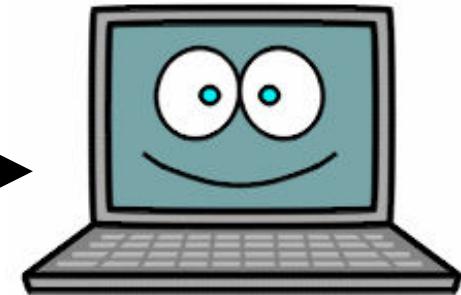
## Adversarial Evaluation: "Adversarial Success"

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

How are you?



I think "P= 90%  
human generated"

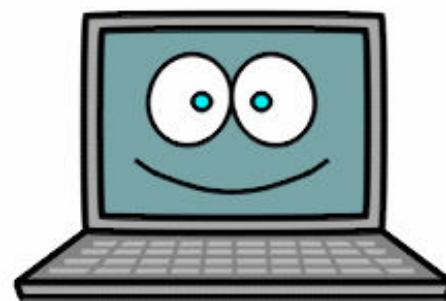


The machine  
evaluator is fooled!!!

# Results: Adversarial Learning Improves Response Generation



Human Evaluator



Machine Evaluator

vs a vanilla generation model

Adversarial Win	Adversarial Lose	Tie
62%	18%	20%

Adversarial Success  
(How often can you fool a machine)

Adversarial Learning	8.0
Standard Seq2Seq model	4.9

# Adversarial Learning for Neural Dialogue Generation

**For** number of training iterations **do**

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

**Update the Discriminator**

**For**  $i=1, G$ -steps **do**

- Sample  $(X, Y)$  from real data
- Sample  $\hat{Y} \sim G(\cdot | X)$
- Compute Reward  $r$  for  $(X, \hat{Y})$  using  $D$ .
- Update  $G$  on  $(X, \hat{Y})$  using reward  $r$

Teacher-Forcing: Update  $G$  on  $(X, Y)$

**End**

**End**

**Update the Generator**

The discriminator forces the generator to produce correct responses

# Human Evaluation



Setting	adver-win	adver-lose	tie
single-turn	0.62	0.18	0.20
multi-turn	0.72	0.10	0.18

The previous RL model only perform better on multi-turn conversations

# Sample response

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

System	Response
Vanilla-MLE	I'm not a doctor.
Vanilla-Sample	Well everything you did was totally untrue.
REINFORCE	I don't know how long it's been.
REGS Monte Carlo (Reward for Every Generation Step)	A few months, I guess.

# Outline

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

# Introduction

Case 1



How do you like CS224s?

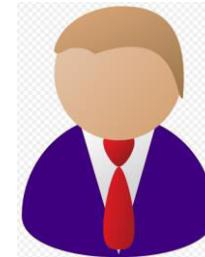
# Introduction

Case 1



How do you like CS224s?

What is CS224n?



# Introduction

Case 1



How do you like CS224s?

What is CS224n?

Is it a class about speech?



# Introduction

Case 1



How do you like CS224s?



CS224n is a class about speech and dialogue

What is cs224n?



# Introduction

Case 1

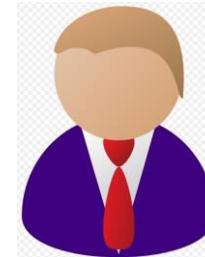


How do you like CS224s?

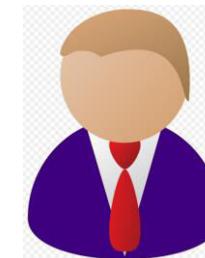


CS224n is a class about speech and dialogue

Who is Hom Tanks?



Oh. Yeah. I like the class.



# Introduction



How do you like CS224s?

What will Current Chatbot  
Systems Do ?

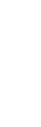
# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

How do you like CS224n?



UNK



# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

How do you like **UNK** ?



# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

How do you like **UNK** ?



Give an output anyway



# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

How do you like **UNK** ?



Forward  
Backward

softmax



# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

How do you like UNK ?



Forward  
Backward

softmax



output

I hate it. It's really aweful



# Introduction

What will Current Chatbot Systems Do ?



How do you like CS224s?

Searching the Web for "how do you like  
CS224n"



We need to the bot to ask questions !!

We Identify three classes of Questions :

three classes of Questions :

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

Which movie did Tom Hanks star in ?

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

three classes of Questions :

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

# Text Clarification

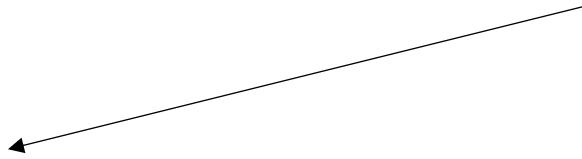
# Text Clarification



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



**What do you mean?**



# Text Clarification



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

AQ



**What do you mean?**



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

# Text Clarification



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



**What do you mean?**



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



**Forest Gump.**



# Text Clarification



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



What do you mean?



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+)

# Text Clarification



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



**What do you mean?**



**Do you mean which movie  
Tom Hanks appear in ?**



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



**Forest Gump.**



**That's correct (+)**

# Text Clarification



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

AQ

QA



**What do you mean?**



**Do you mean which movie  
Tom Hanks appear in ?**



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



**That's correct (+)**



**Forest Gump.**



**That's correct (+)**

# Text Clarification



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

AQ

QA



**What do you mean?**



**Do you mean which movie  
Did Tom Hanks appear in ?**



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



**That's correct (+)**



**Forest Gump.**



**Forest Gump.**



**That's correct (+)**

# Text Clarification



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

AQ



What do you mean?

QA



Do you mean which movie  
Tom Hanks appear in ?



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+)



That's correct (+)



Forest Gump.



That's correct (+)

three classes of Questions :

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

## Case 2: Knowledge Operation.

Knowledge  
Base:

**Tom Hanks directed Larry Crowne**

**Tom Hanks starred Forrest Gump**

**Robert Zemeckis directed Forrest Gump**

**which movie Did Tom Hanks appear in ?**

## Case 2: Knowledge Operation.

Knowledge  
Base:

Tom Hanks directed Larry Crowne

Tom Hanks starred Forrest Gump

Robert Zemeckis directed Forrest Gump

which movie Did Tom Hanks appear in ?

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : Can you give me a hint ?

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : Can you give me a hint ?



T : It is related to the kb fact that

Forrest Gump starred\_actors Tom Hanks

Robin Wright, Gary Sinise

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : Can you give me a hint ?



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



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

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : Can you give me a hint ?



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



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



T : That's correct. (+)

## Case 2: Knowledge Operation.

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks,  
Robin Wright, Gary Sinise

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : Can you give me a hint ?



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



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



T : That's correct. (+)



S : Forrest Gump



S : Forrest Gump

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?

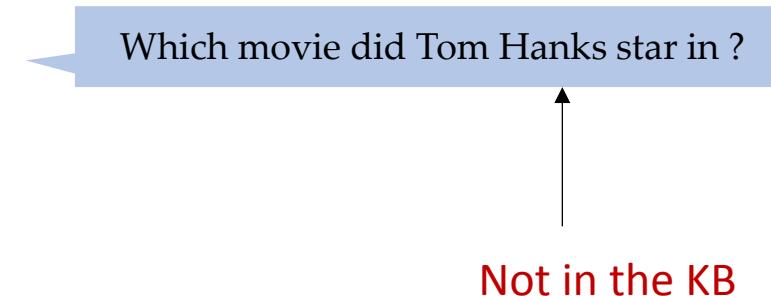
## three classes of Questions :

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

How do you like Hom Tanks?

## three classes of Questions :

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



## Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?

Not in the KB

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?



T : The answer is Forrest Gump.

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?



T : The answer is Forrest Gump.

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?



T : The answer is Forrest Gump.

... Other questions/ Other answers

In what scenarios does a bot  
need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?



T : The answer is Forrest Gump.

... Other questions/ Other answers



T : Which movie did Tom Hanks star in ?

In what scenarios does a bot need to ask questions ?

### Case 3: Knowledge Acquisition .

kb: Larry Crowne directed\_by Tom Hanks

kb: Forrest Gump starred\_actors Tom Hanks

kb: Forrest Gump starred\_actors Sally Field

kb: Forrest Gump directed\_by Robert Zemeckis



T : Which movie did Tom Hanks star in ?



S : I don't know. What's the answer?



T : The answer is Forrest Gump.

... Other questions/ Other answers



T : Which movie did Tom Hanks star in ?

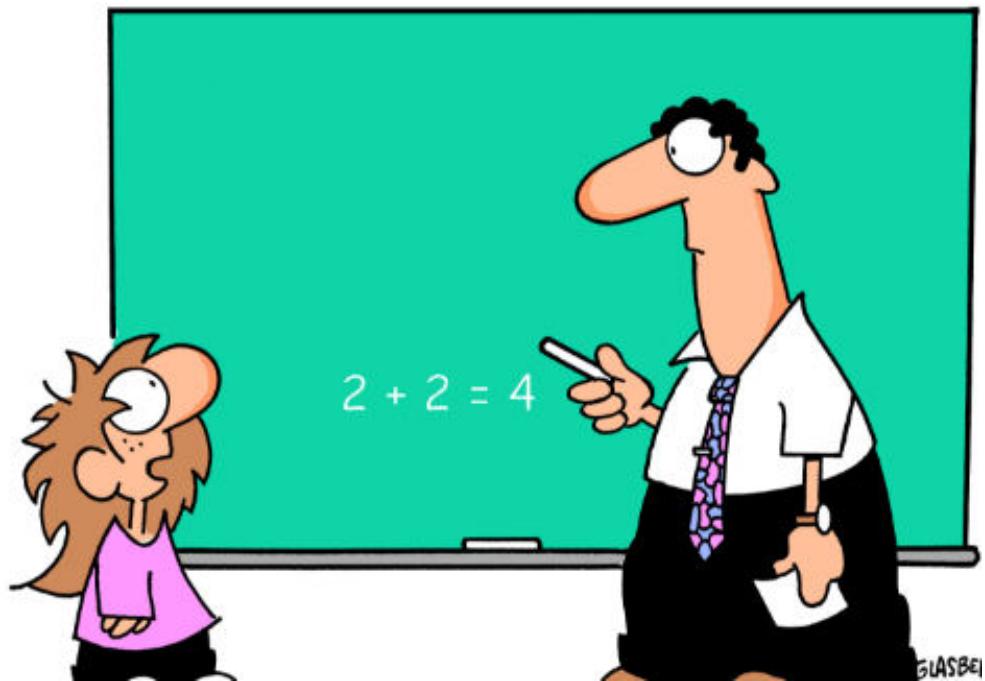


S : Forrest Gump

three classes of Questions :

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

When shall a bot asks a question ?



Shall I ask a question ???



Which movvie did Tom Hanks sttar in?



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?  
-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Forest Gump.



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+) +1



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+) +1

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Larry Crowne



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Larry Crowne



That's incorrect (+) -1



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Larry Crowne



That's incorrect (+) -1

Reward: -1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ

QA



What do you mean?

-CostAQ



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+) +1

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ

QA



What do you mean?  
-CostAQ



Larry Crowne



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+) +1

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

QA



Larry Crowne



I mean which film did Tom Hanks appear in.



That's incorrect (--)



Forest Gump.



That's correct (+)

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

QA



Larry Crowne



I mean which film did Tom Hanks appear in.



That's incorrect (--)



Forest Gump.



That's correct (+)

Reward: -1

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

QA



Forest Gump.



I mean which film did Tom Hanks appear in.



Forest Gump.



That's correct (+)

Reward: 1-CostAQ



Which movvie did Tom Hanks sttar in?

AQ



What do you mean?

QA



Forest Gump.



I mean which film did Tom Hanks appear in.



That's correct (+)



Forest Gump.



That's correct (+)

Reward: 1-CostAQ



Which movie did Tom Hanks star in?

AQ



What do you mean?

QA



Forest Gump.



I mean which film did Tom Hanks appear in.



That's correct (+)



Forest Gump.



That's correct (+)

Reward: +1

Reward: 1-CostAQ

Final Answer Correct  
Final Answer InCorrect

Asking Question  
 $1 - cost_{AQ}$   
 $-1 - cost_{AQ}$

Not asking Question  
1  
-1

## Setting2: Reinforcement Learning



**Which movvie did Tom Hanks sttar in?**



**Ask a question or not**

**Policy Gradient**

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

## Setting2: Reinforcement Learning



T : Which movie did Tom Hanks star in ?



Ask a question or not .....

Memory Network

	Asking Question	Not asking Question
Final Answer Correct	$1 - cost_{AQ}$	1
Final Answer InCorrect	$-1 - cost_{AQ}$	-1

Policy Gradient

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

## Setting2: Reinforcement Learning



T : Which movie did Tom Hanks star in ?



Ask a question or not .....

Memory Network

	Asking Question	Not asking Question
Final Answer Correct	$1 - cost_{AQ}$	1
Final Answer InCorrect	$-1 - cost_{AQ}$	-1

Policy Gradient

$$\nabla J(\theta) = \nabla \log p_{RL}(a)[r - b] \leftarrow \text{Baseline}$$

# Q&A