Data-Driven Dialogue

Alan Ritter
Computer Science and Engineering
The Ohio State University

Collaborators:
Jiwei Li, Dan Jurafsky (Stanford)
Colin Cherry (Google)
Bill Dolan, Michel Galley, Jianfeng Gao (Microsoft Research),
Ashutosh Baheti (Ohio State)

THE OHIO STATE UNIVERSITY

1950s ~ 2010

Dialog systems mostly rule-based

Chatbots:

Rule-Based: Eliza (Weizenbaum 1966)

Information Retrieval (Isbell et. al. 2000)

Goal-Directed Dialogue Systems:

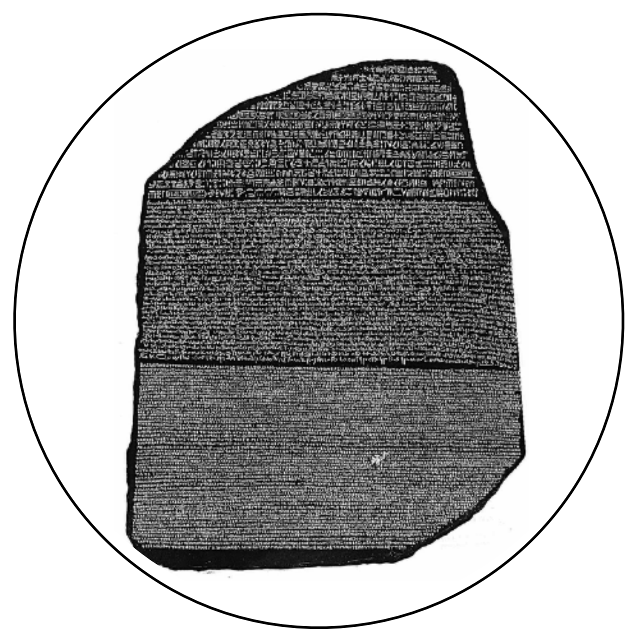
ATIS Dataset (Hemphill, 1990)

-Manually annotated

1990s ~ 2010s

Data-Driven Machine Translation

Learning from millions of bilingual documents on the web



2011 ~ Today

Data-Driven Dialogue
500 million conversations per month on Twitter alone

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

2011 ~ Today

Data-Driven Dialogue

500 million conversations per month on Twitter alone

(vs. 30m for French-English translation)

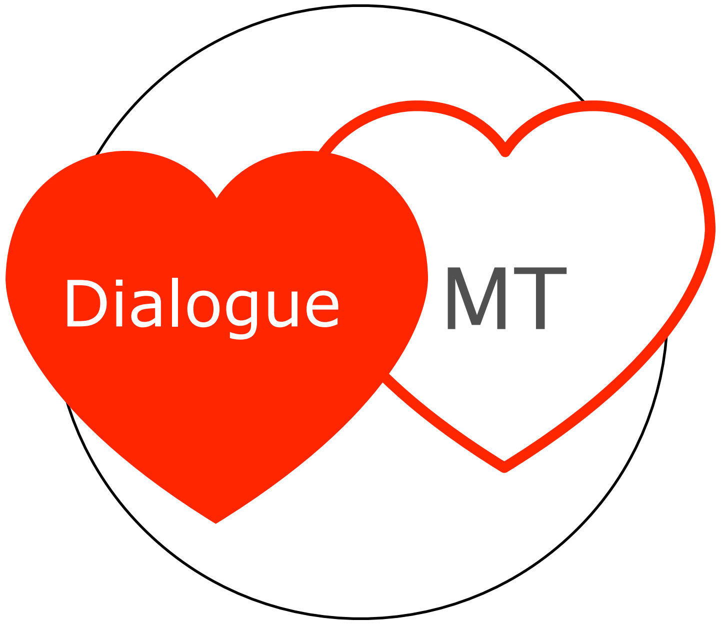
Alan Ritter, Colin Cherry, Bill Dolan (EMNLP 2011) “Data-Driven Response Generation in Social Media”

2011 ~ Today

Data-Driven Dialogue

500 million conversations per month on Twitter alone

(vs. 30m for French-English translation)





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

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

Output:

Yum ! I

Noisy Channel Model

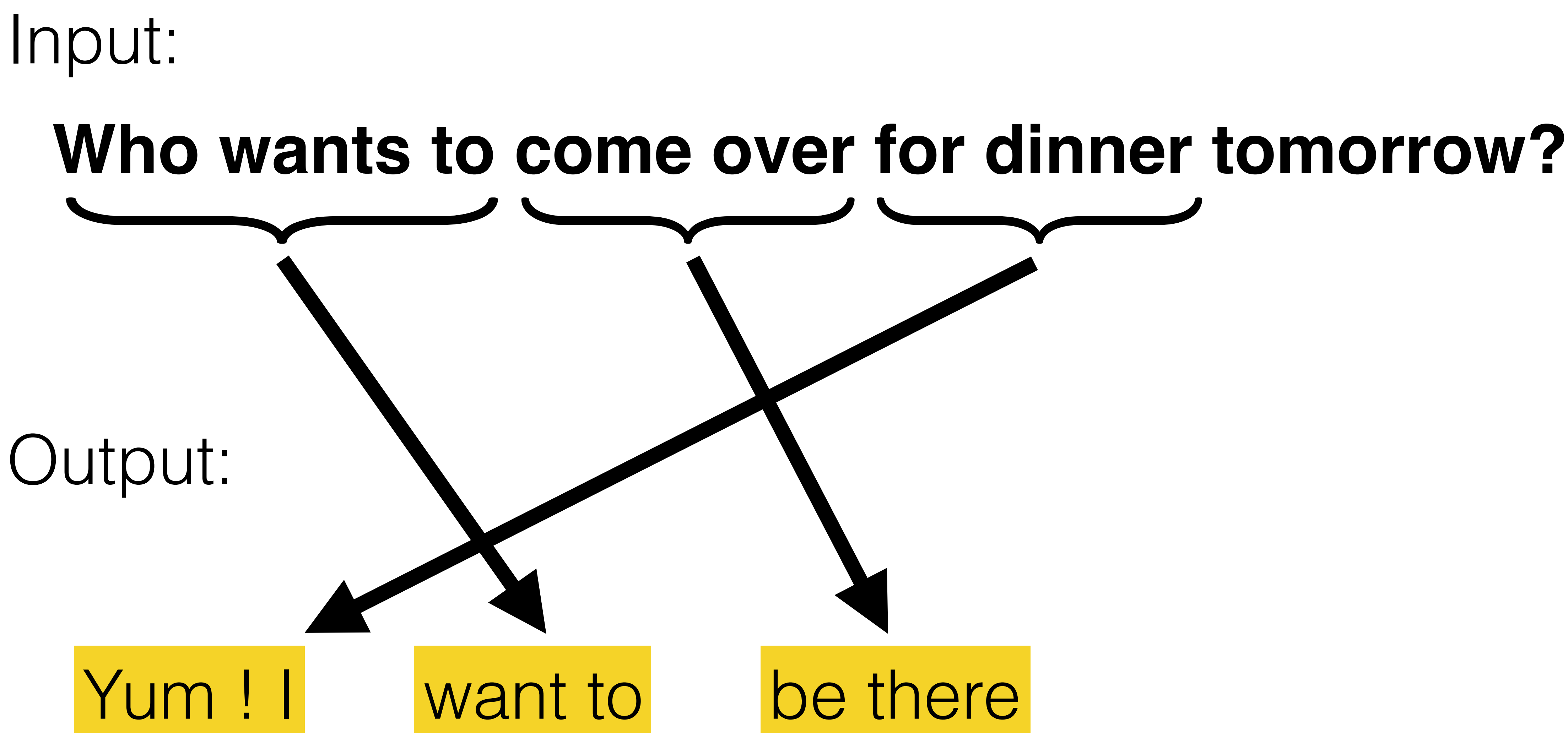
Input:

Who wants to come over for dinner tomorrow?

Output:

Yum ! I want to

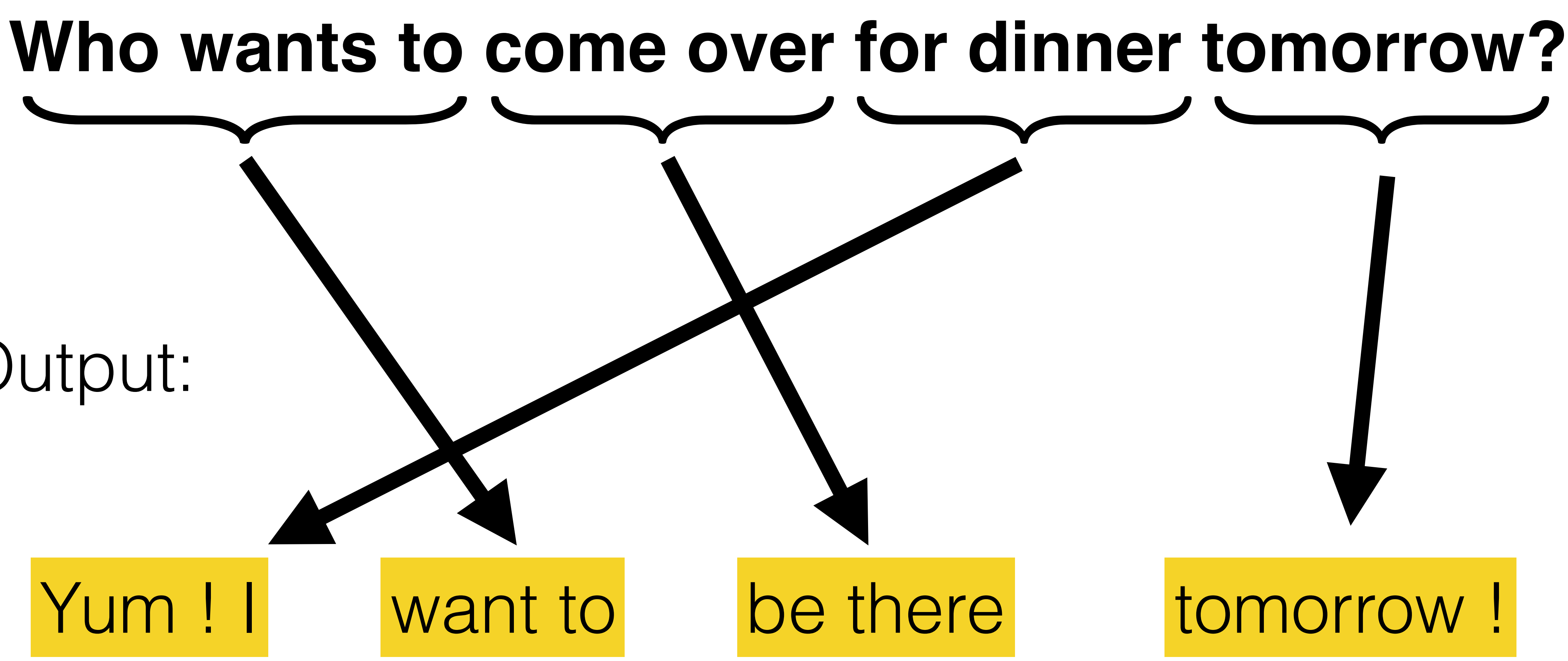
Noisy Channel Model



Noisy Channel Model

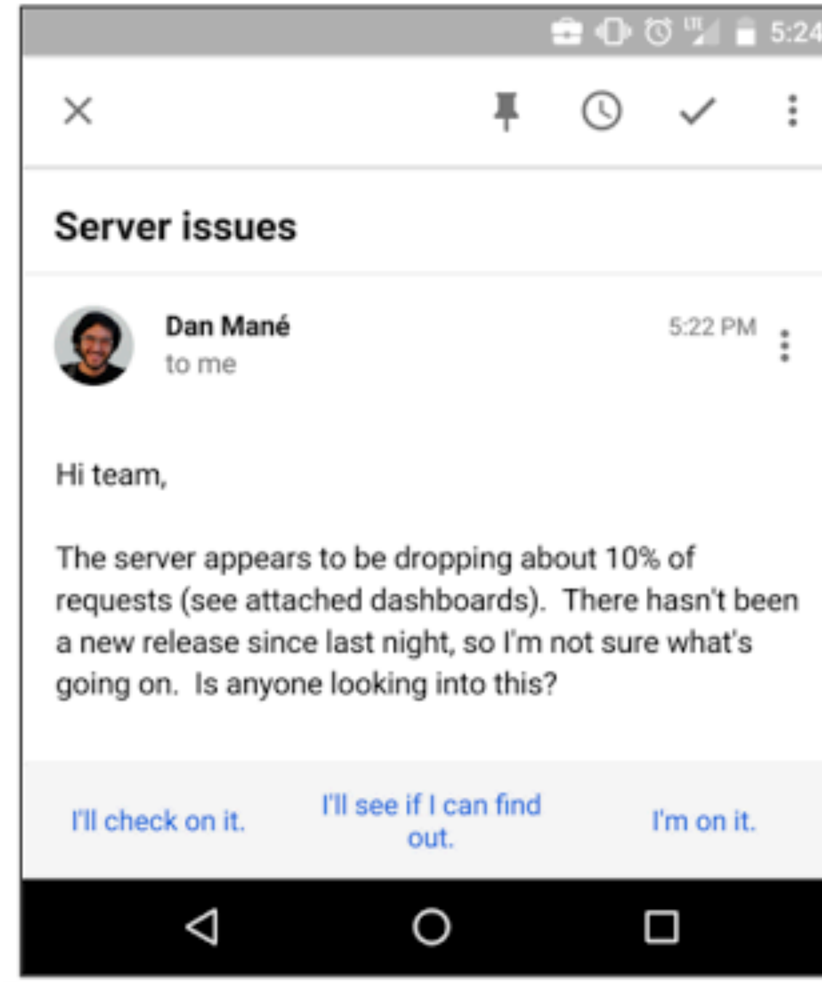
Input:

Output:



Neural Conversation

[Sordoni et. al. 2015] [Xu et. al. 2016] [Wen et. al. 2016]
[Li et. al. 2016] [Kannan et. al. 2016] [Serban et. al. 2016]



Google Research Blog

Computer, respond to this email.

Tuesday, November 03, 2015

Posted by Greg Corrado*, Senior Research Scientist

Another bizarre feature of our early prototype was its propensity to respond with “I love you” to seemingly anything. As adorable as this sounds, it wasn’t really what we were hoping for. Some analysis revealed that the system was doing exactly what we’d trained it to do, generate likely responses -- and it turns out that responses like “Thanks", "Sounds good", and “I love you” are super common -- so the system would lean on them as a safe bet if it was unsure. Normalizing the

Neural Conversation

[Sordoni et. al. 2015] [Xu et. al. 2016] [Wen et. al. 2016]

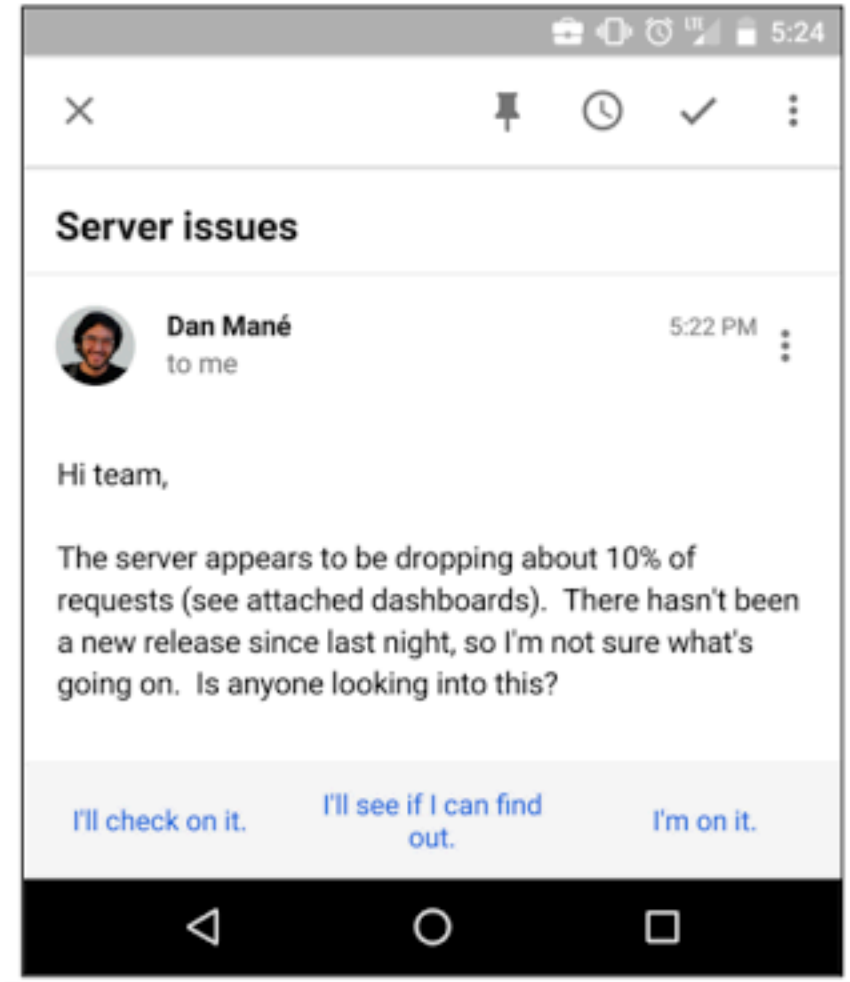
[Li et. al. 2016] [Kannan et. al. 2016] [Serban et. al. 2016]

Google Research Blog

Computer, respond to this email.

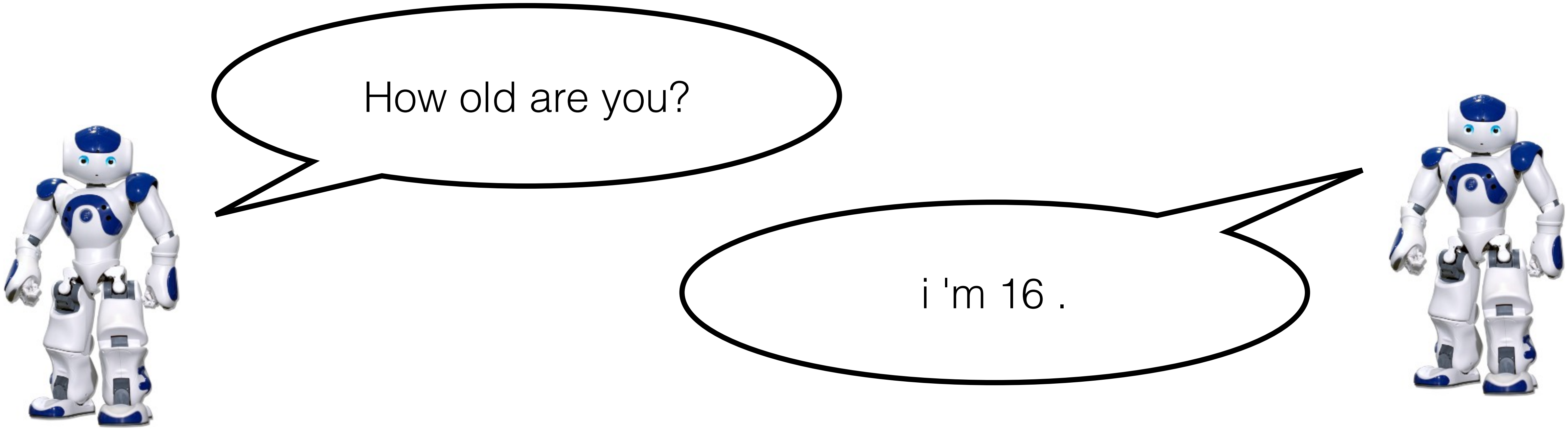
Tuesday, November 03, 2015

Posted by Greg Corrado*, Senior Research Scientist

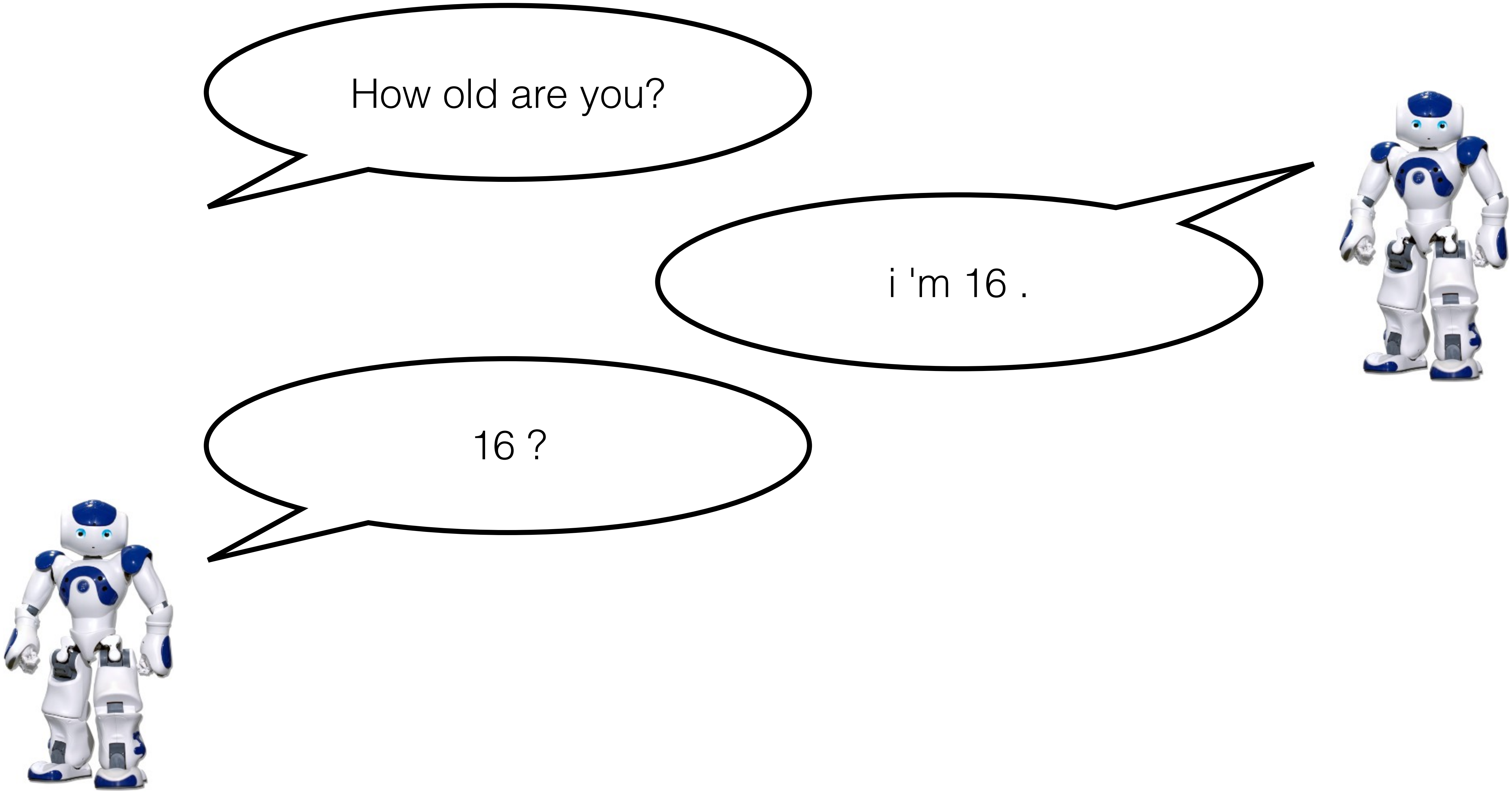


Another bizarre feature of our early prototype was its propensity to respond with "I love you" to seemingly anything. As adorable as this sounds, it wasn't really what we were hoping for. Some analysis revealed that the system was doing exactly what we'd trained it to do, generate likely responses -- and it turns out that responses like "Thanks", "Sounds good", and "I love you" are super common -- so the system would lean on them as a safe bet if it was unsure. Normalizing the

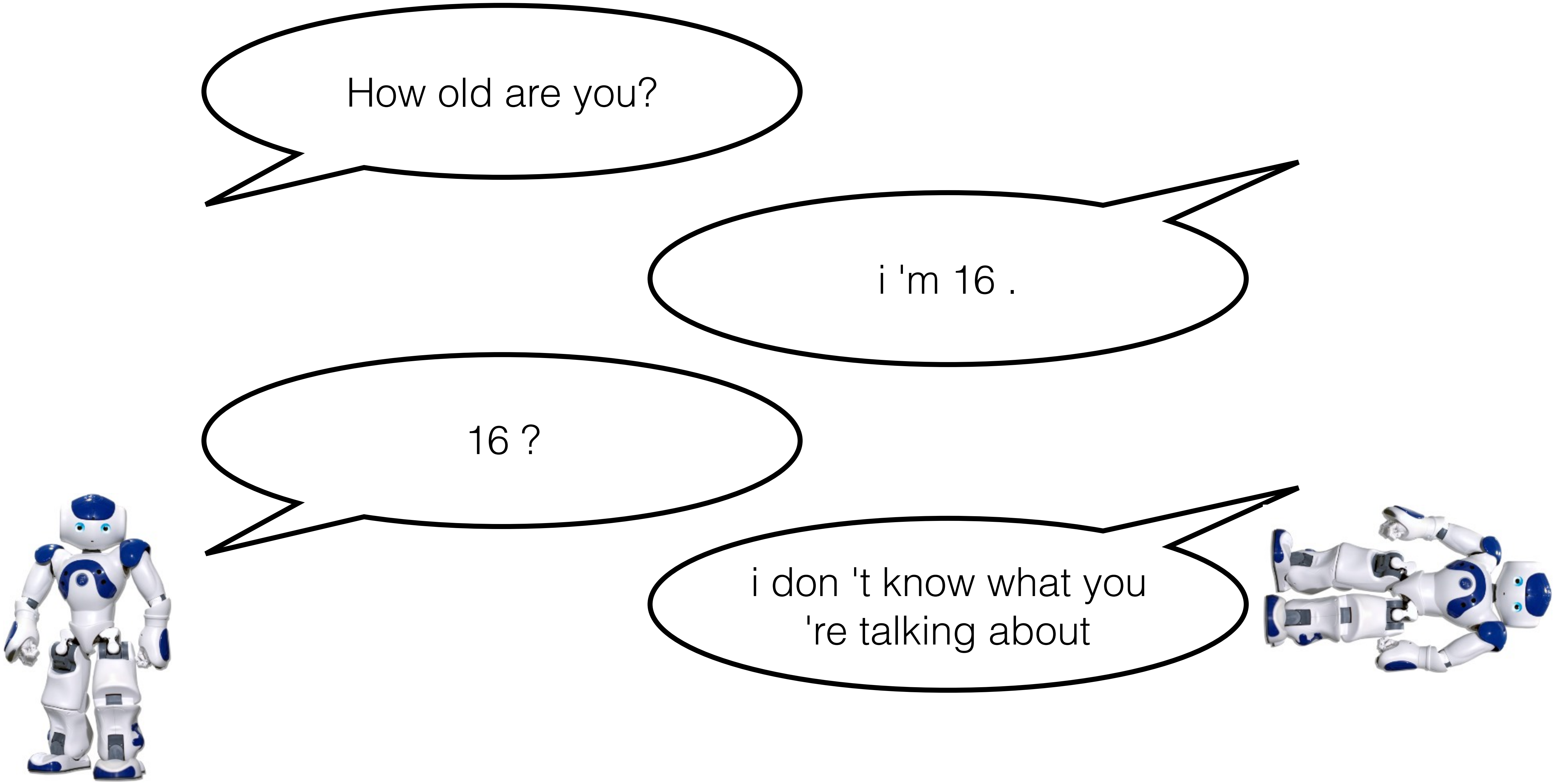
Slide Credit: Jiwei Li



Slide Credit: Jiwei Li



Slide Credit: Jiwei Li



Slide Credit: Jiwei Li

How old are you?

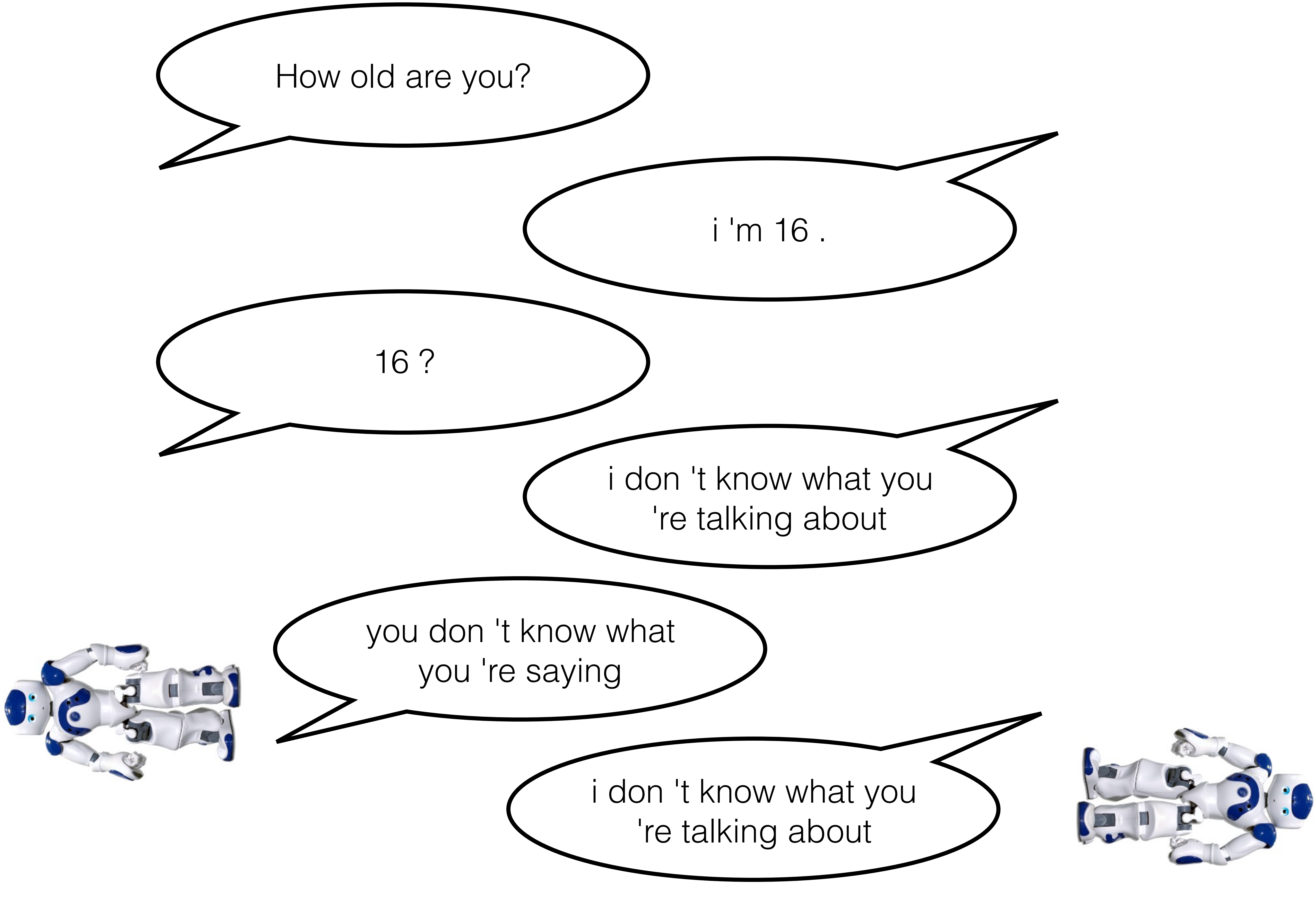
i 'm 16 .

16 ?

i don 't know what you 're talking about

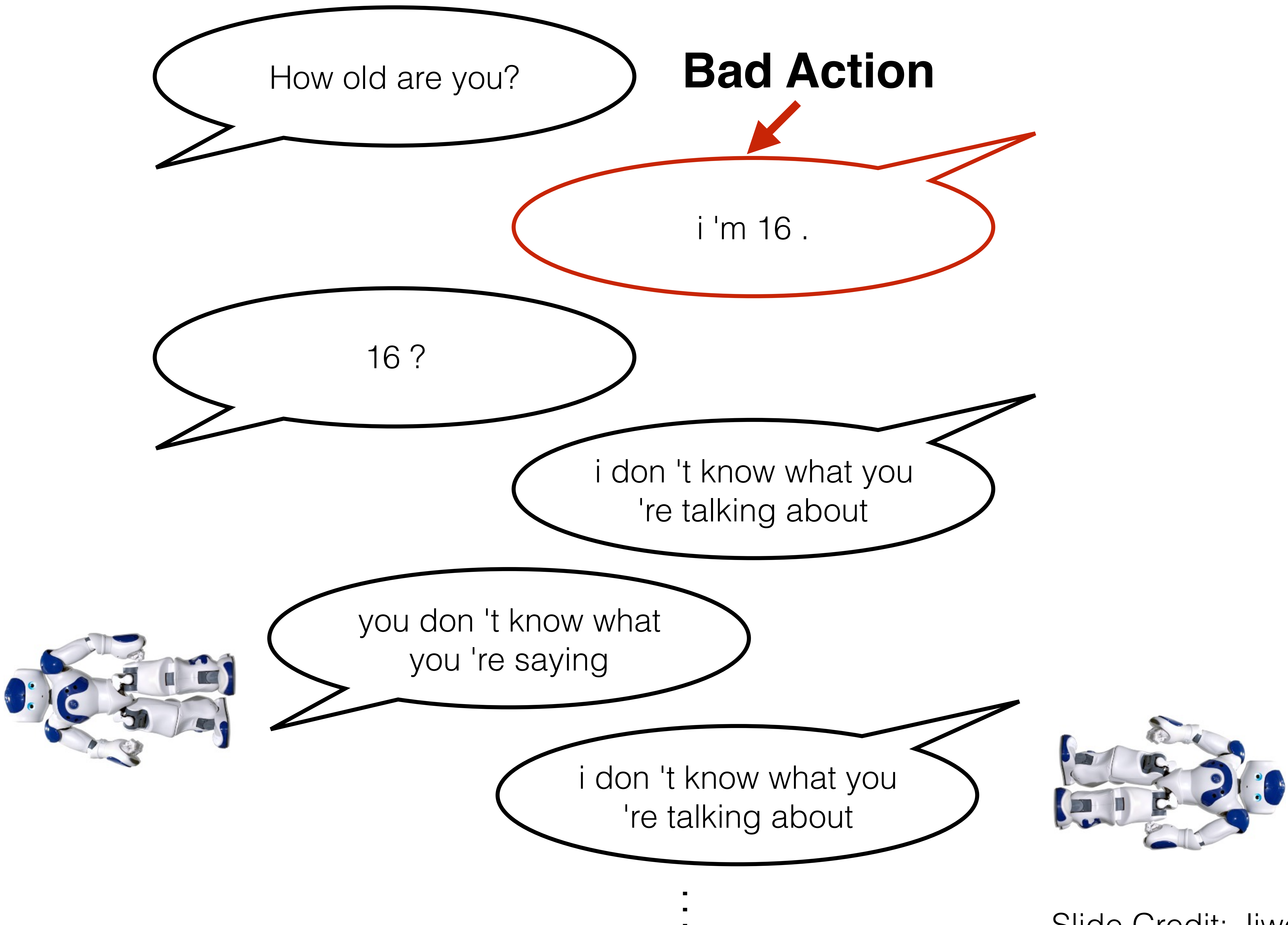
you don 't know what you 're saying

Slide Credit: Jiwei Li

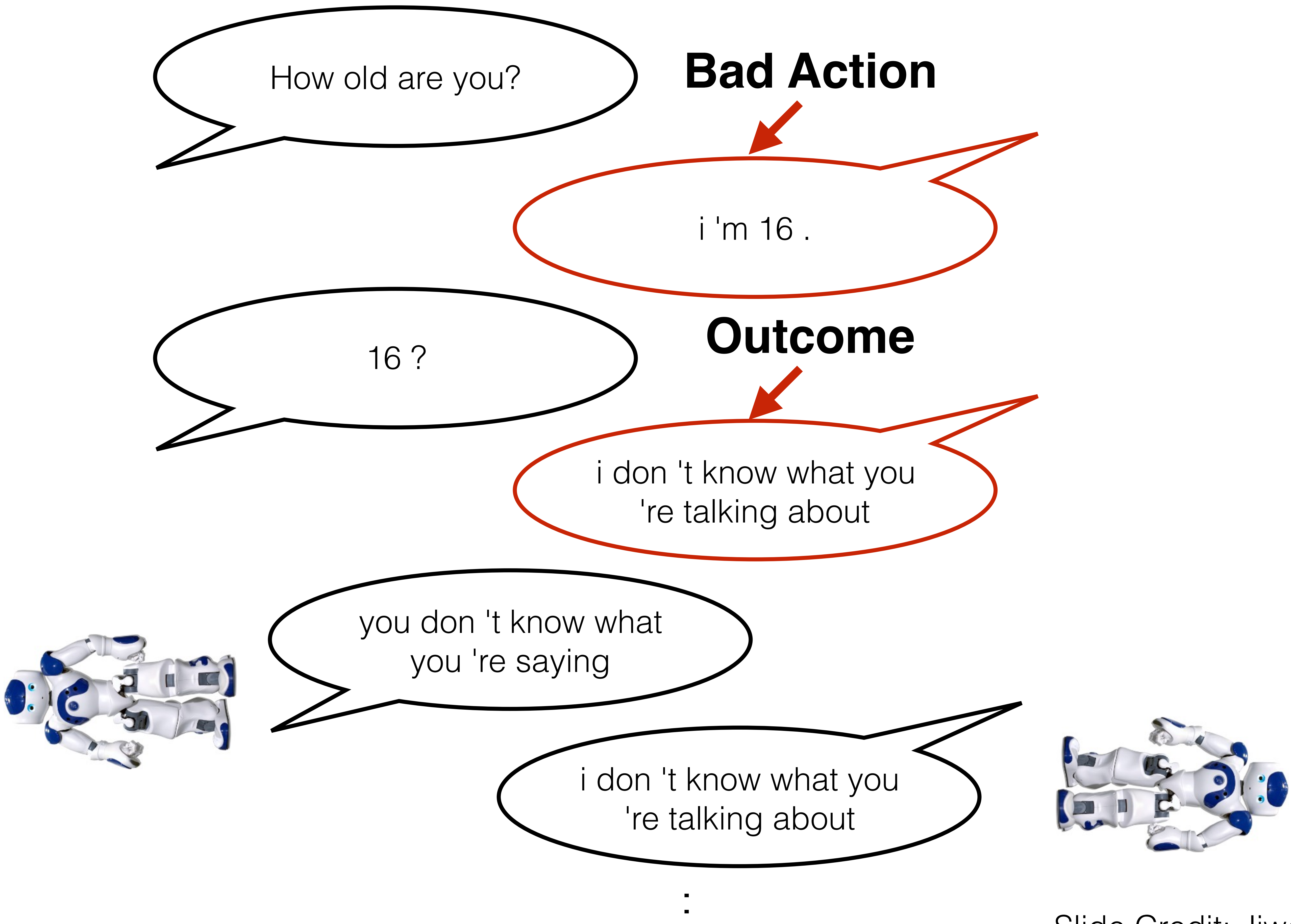


⋮

Slide Credit: Jiwei Li



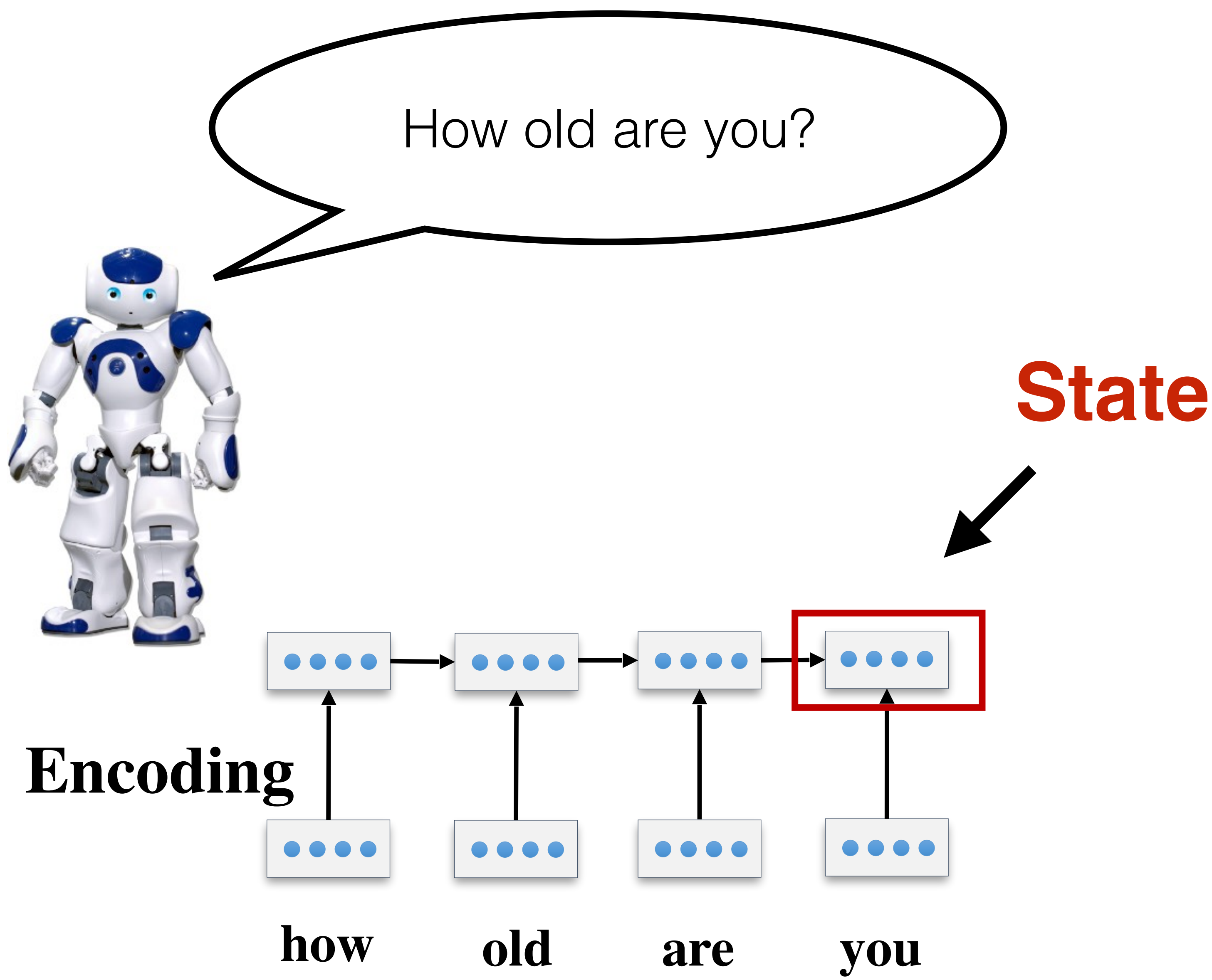
Slide Credit: Jiwei Li



Slide Credit: Jiwei Li

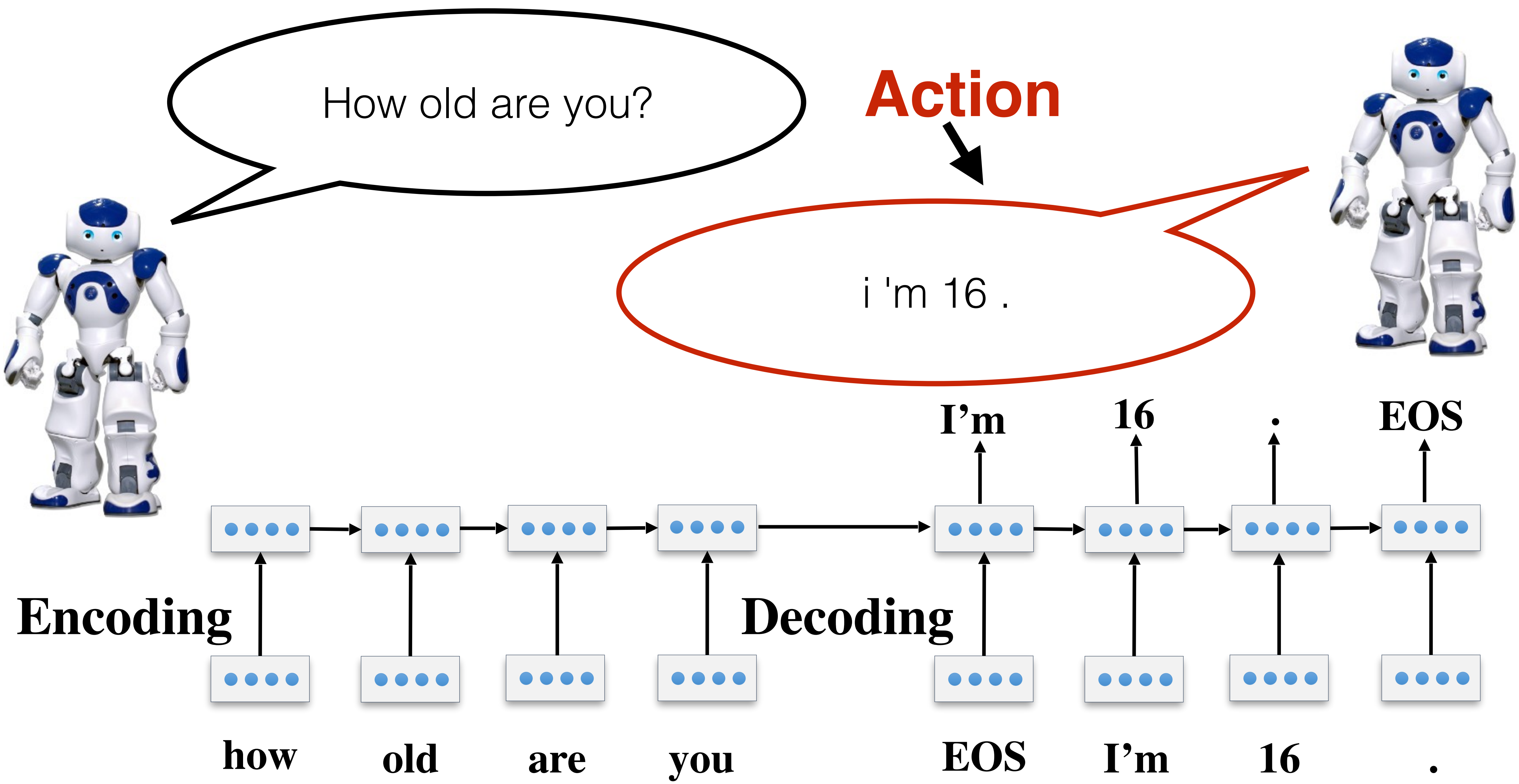
Deep Reinforcement Learning

[Li, Monroe, Ritter, Galley, Gao, Jurafsky EMNLP 2016]



Deep Reinforcement Learning

[Li, Monroe, Ritter, Galley, Gao, Jurafsky EMNLP 2016]



Learning: Policy Gradient

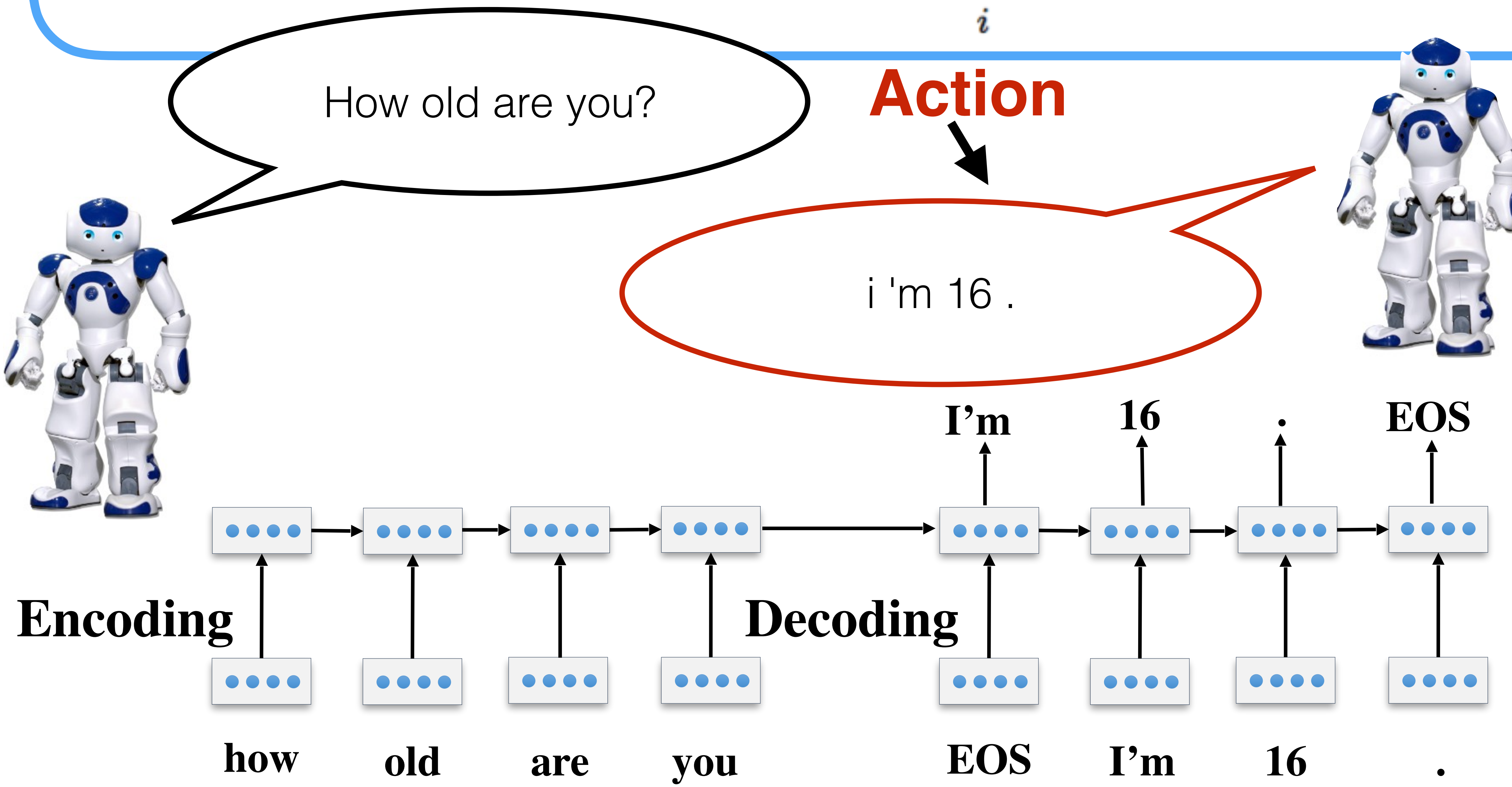
REINFORCE Algorithm (Williams,1992)

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

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

What we want to learn

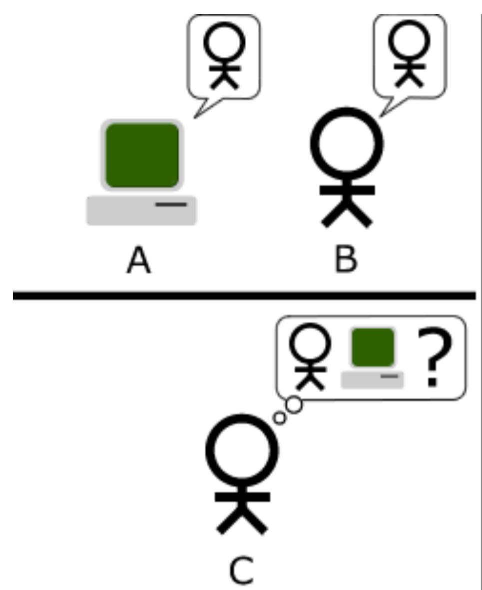
$$\nabla J(\theta) = \nabla \log \prod_i \boxed{p(s_i|s_{i-1})} R(s_1, s_2, ..., s_N)$$

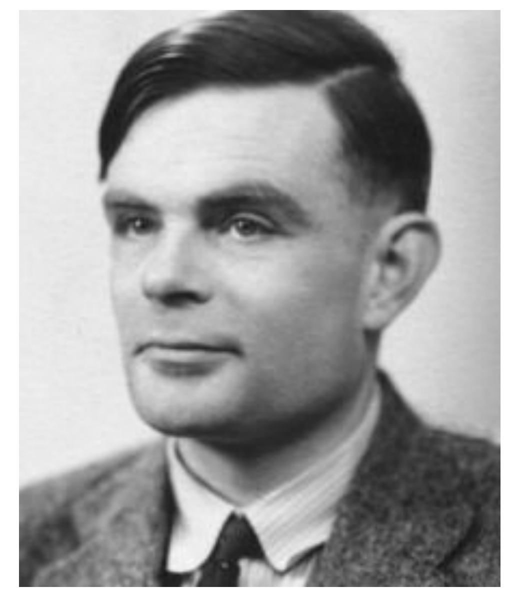


Q: Rewards?

Q: Rewards?

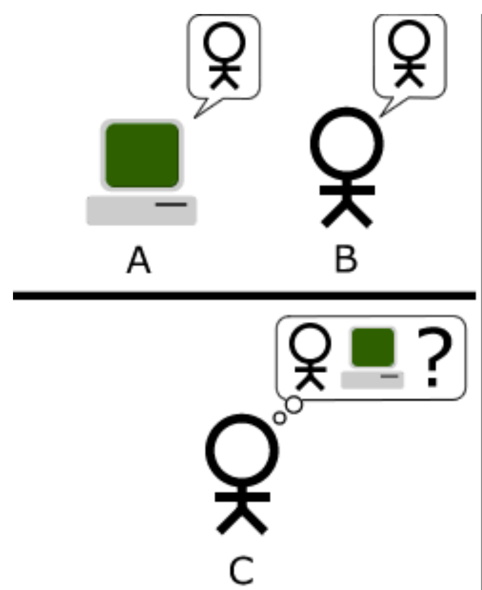
A: Turing Test

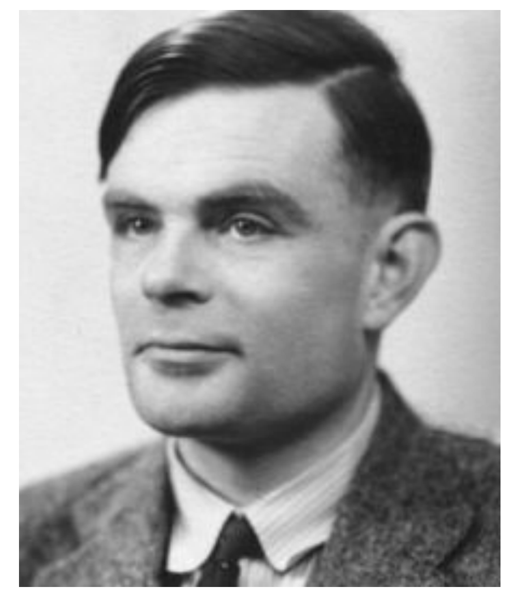




Q: Rewards?

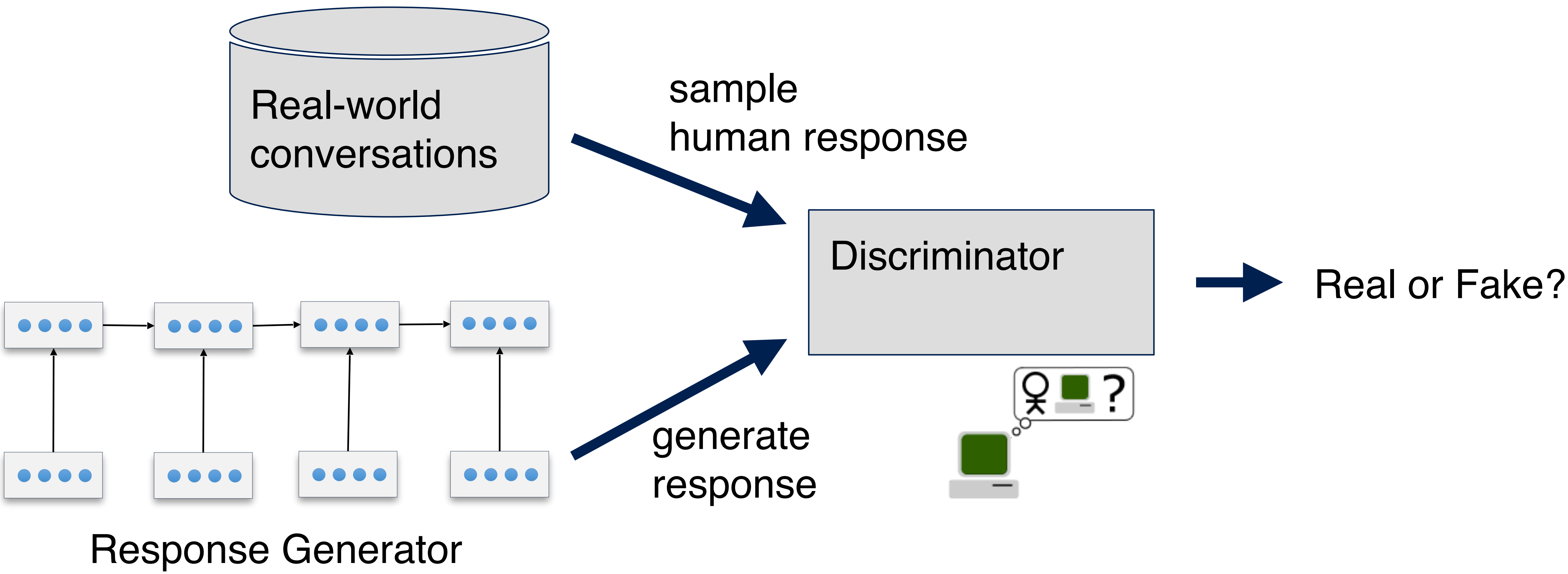
A: Turing Test





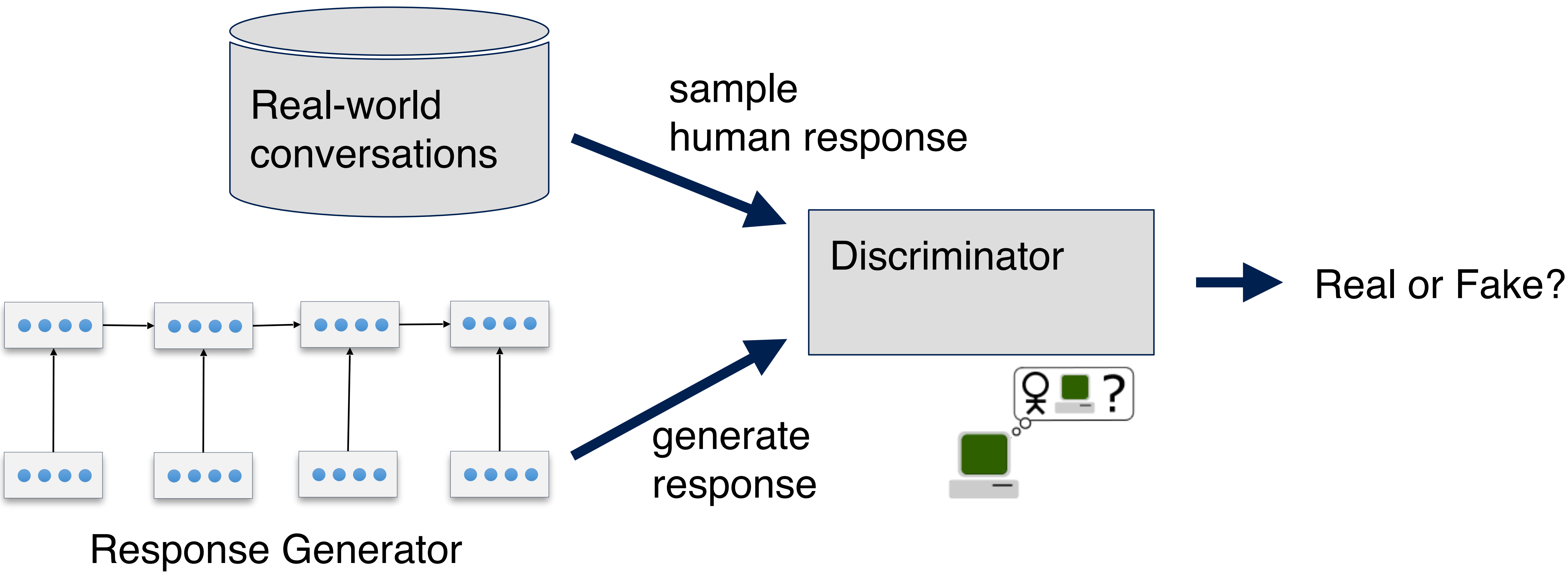
Adversarial Learning
(Goodfellow et al., 2014)

Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2017]



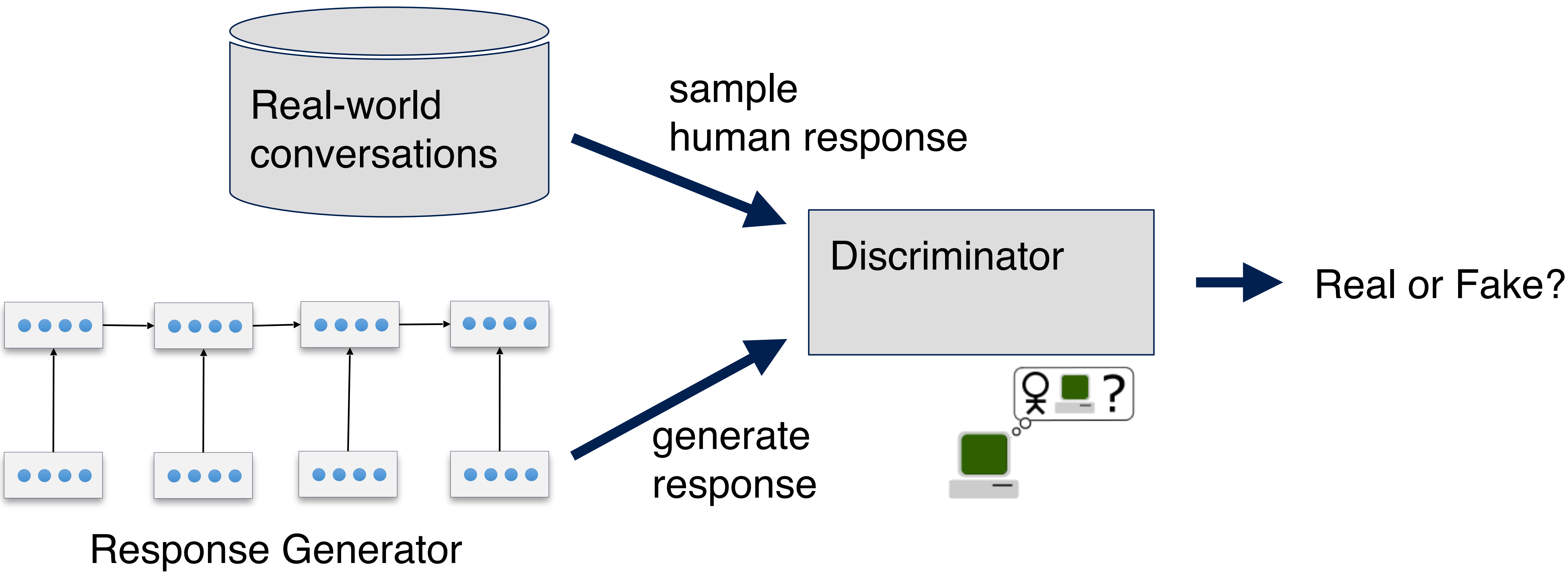
Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2017]

(Alternate Between Training Generator and Discriminator)



Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2017]

(Alternate Between Training Generator and Discriminator)



REINFORCE Algorithm (Williams,1992)

Adversarial Learning Improves Response Generation

Human Evaluator:

vs vanilla generation model

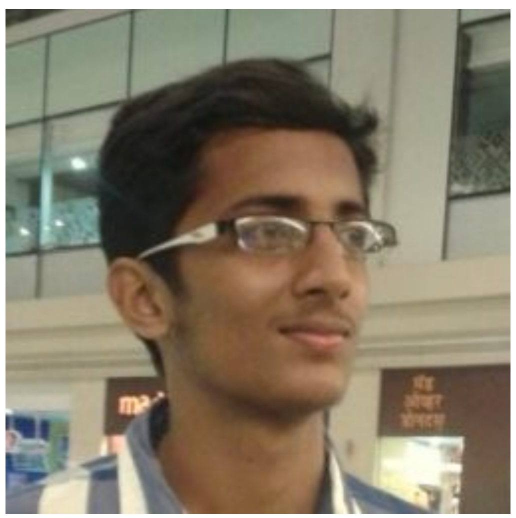
Adversarial Win	Adversarial Lose	Tie
62%	18%	20%

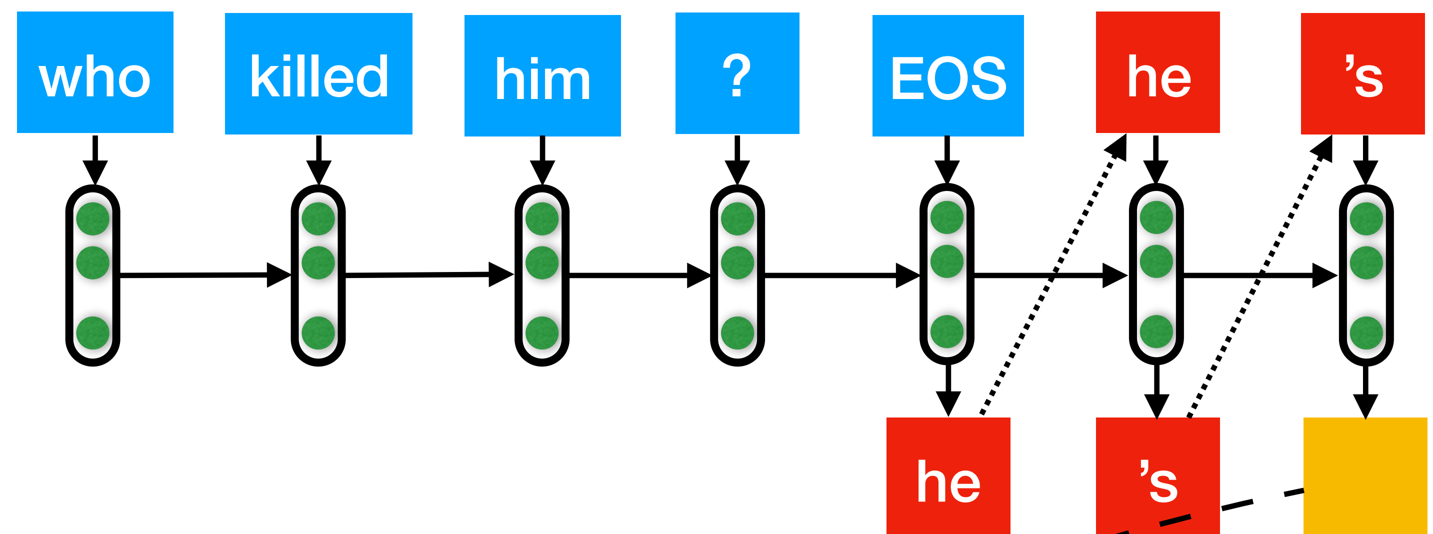
Machine Evaluator:
[Bowman et. al. 2016]

Adversarial Success (How often can you fool a machine)

Adversarial Learning	8.0%
Standard Seq2Seq model	4.9%

Another Approach: Distributional Constraints

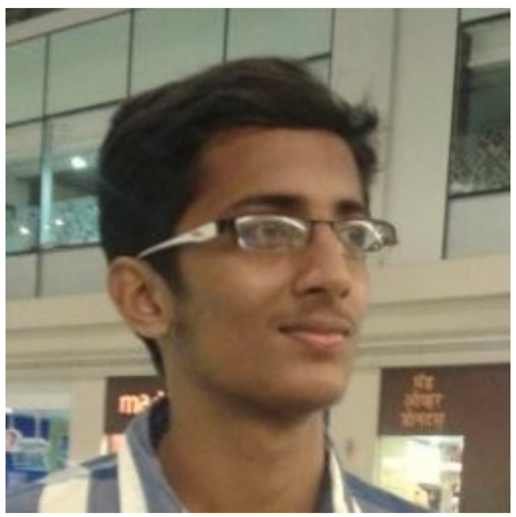




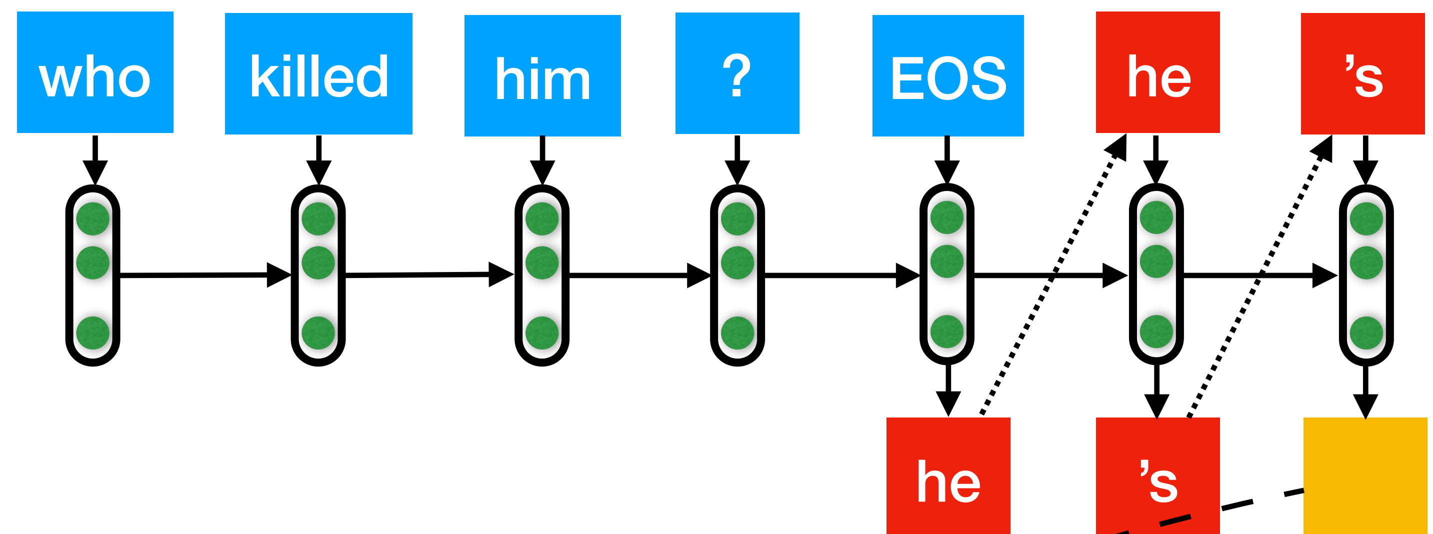
Stop word	Likelihood score
a	-4.62
the	-5.69
in	-5.95
<unk>	-6.26
on	-6.97
an	-7.00
my	-7.31
not	-7.57

Topic word	Likelihood score
shot	-6.58
dead	-6.95
head	-11.67
died	-12.24
murder	-12.43
president	-12.56
evil	-12.66
father	-12.66

Another Approach: Distributional Constraints



See our paper!

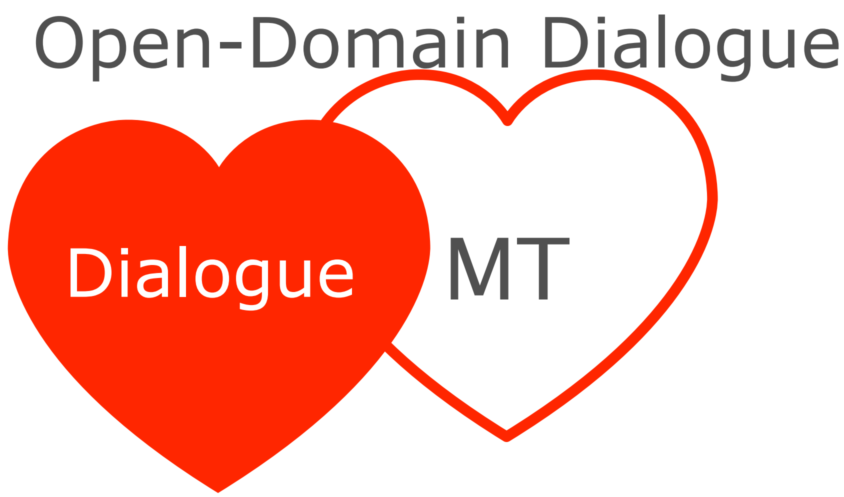


Stop word	Likelihood score
a	-4.62
the	-5.69
in	-5.95
<unk>	-6.26
on	-6.97
an	-7.00
my	-7.31
not	-7.57

Topic word	Likelihood score
shot	-6.58
dead	-6.95
head	-11.67
died	-12.24
murder	-12.43
president	-12.56
evil	-12.66
father	-12.66

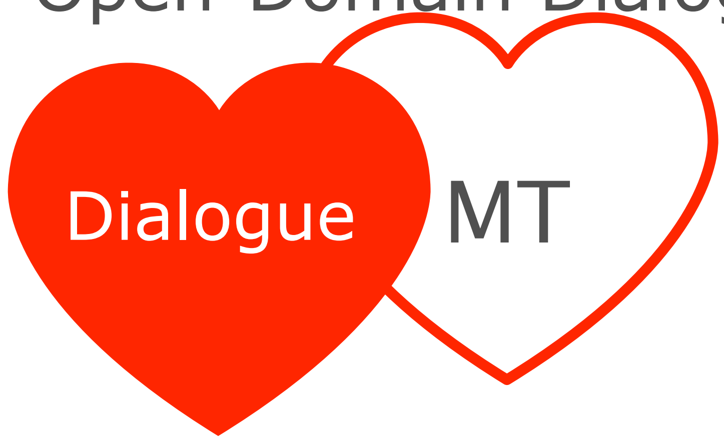
Takeaways

Takeaways

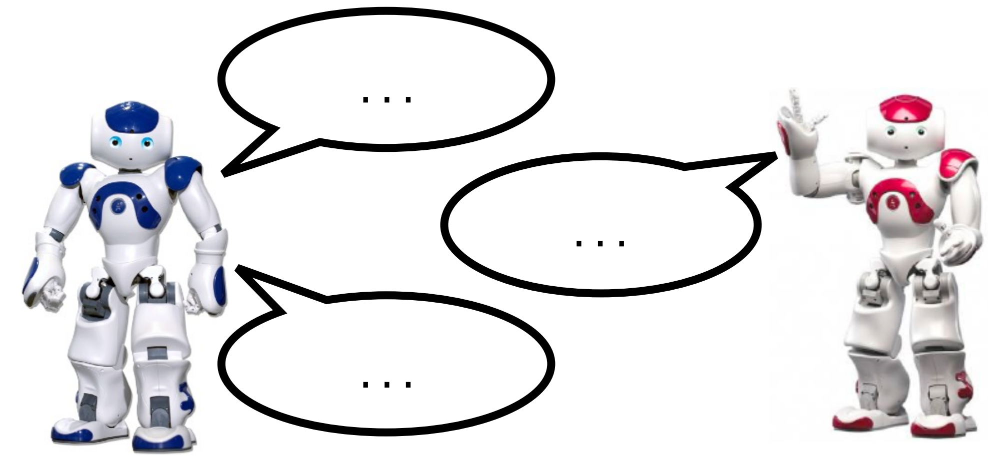


Takeaways

Open-Domain Dialogue

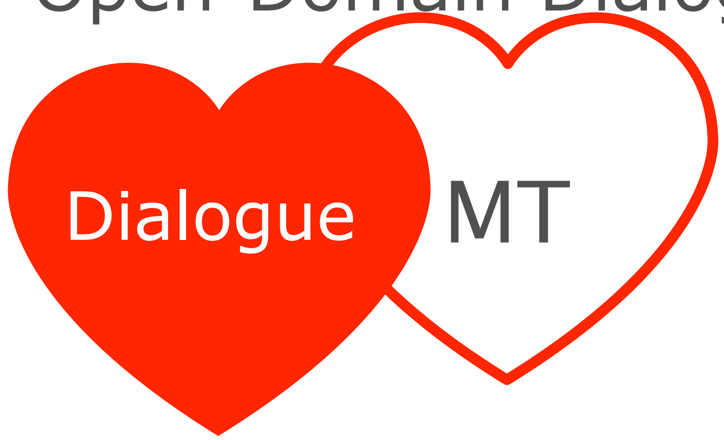


Learning from Delayed-Reward

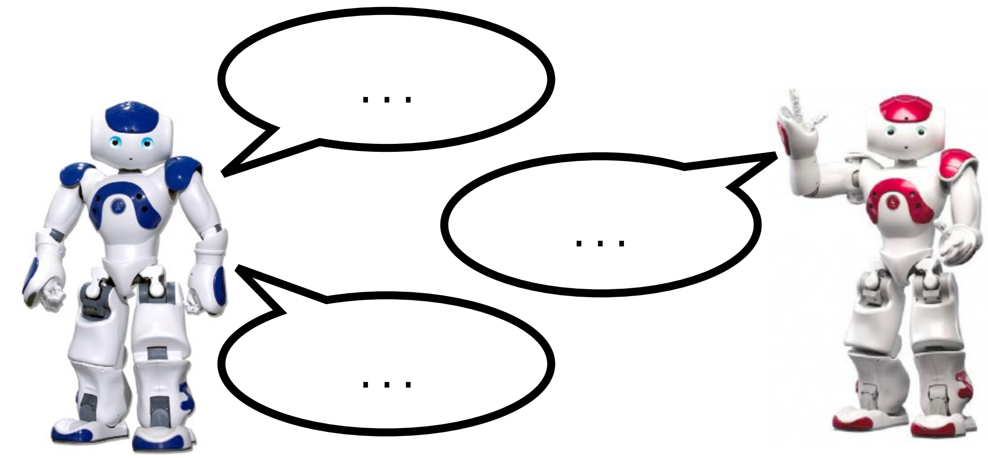


Takeaways

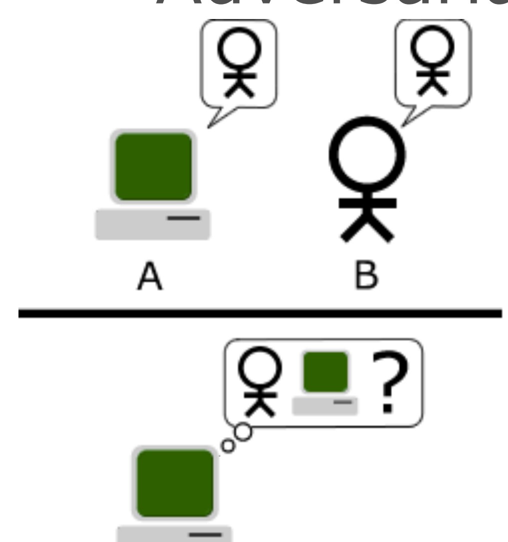
Open-Domain Dialogue

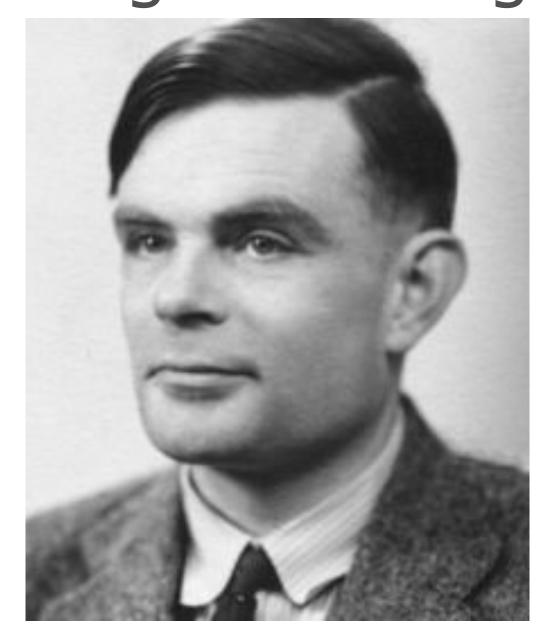


Learning from Delayed-Reward



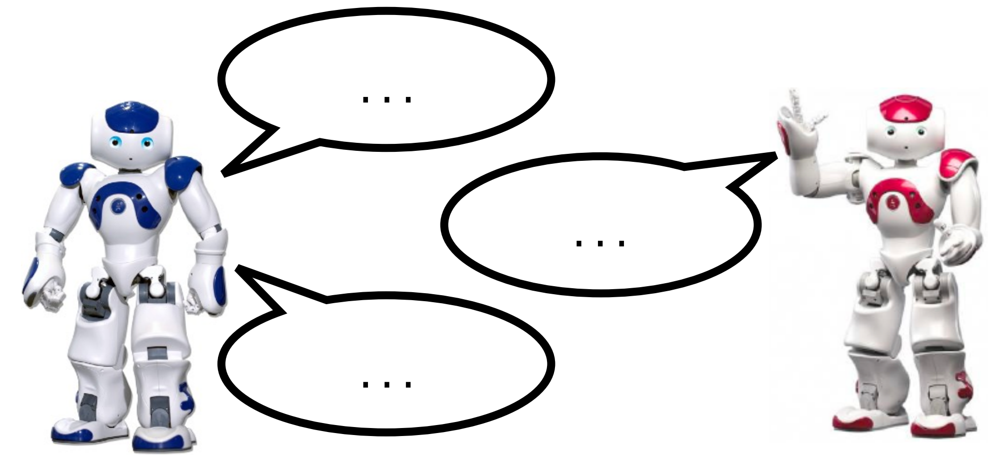
Adversarial Learning for Dialogue



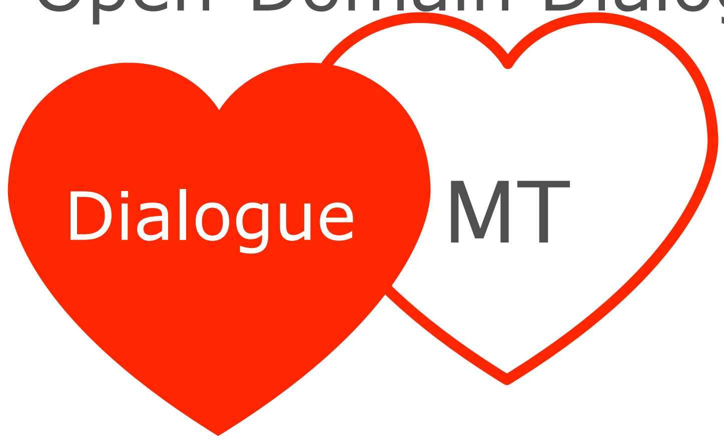


Takeaways

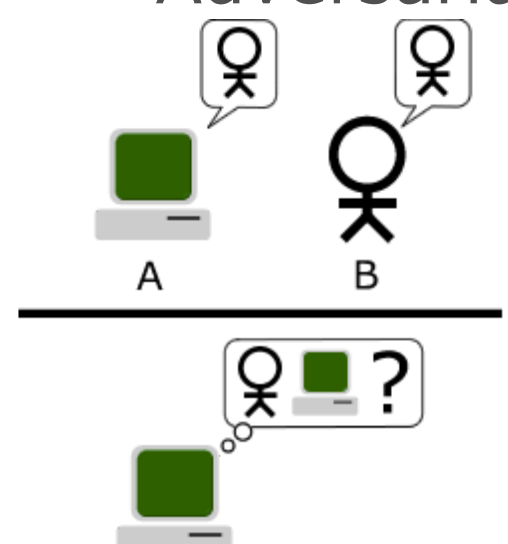
Learning from Delayed-Reward

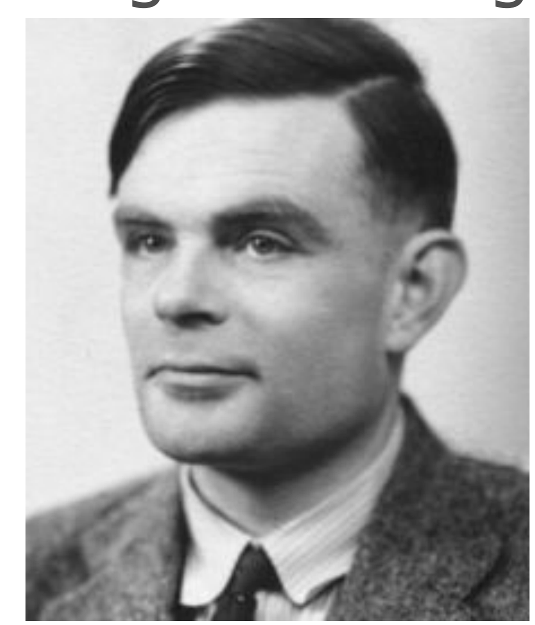


Open-Domain Dialogue



Adversarial Learning for Dialogue





Thank You!