Improving Sequence Generation by GAN

Outline

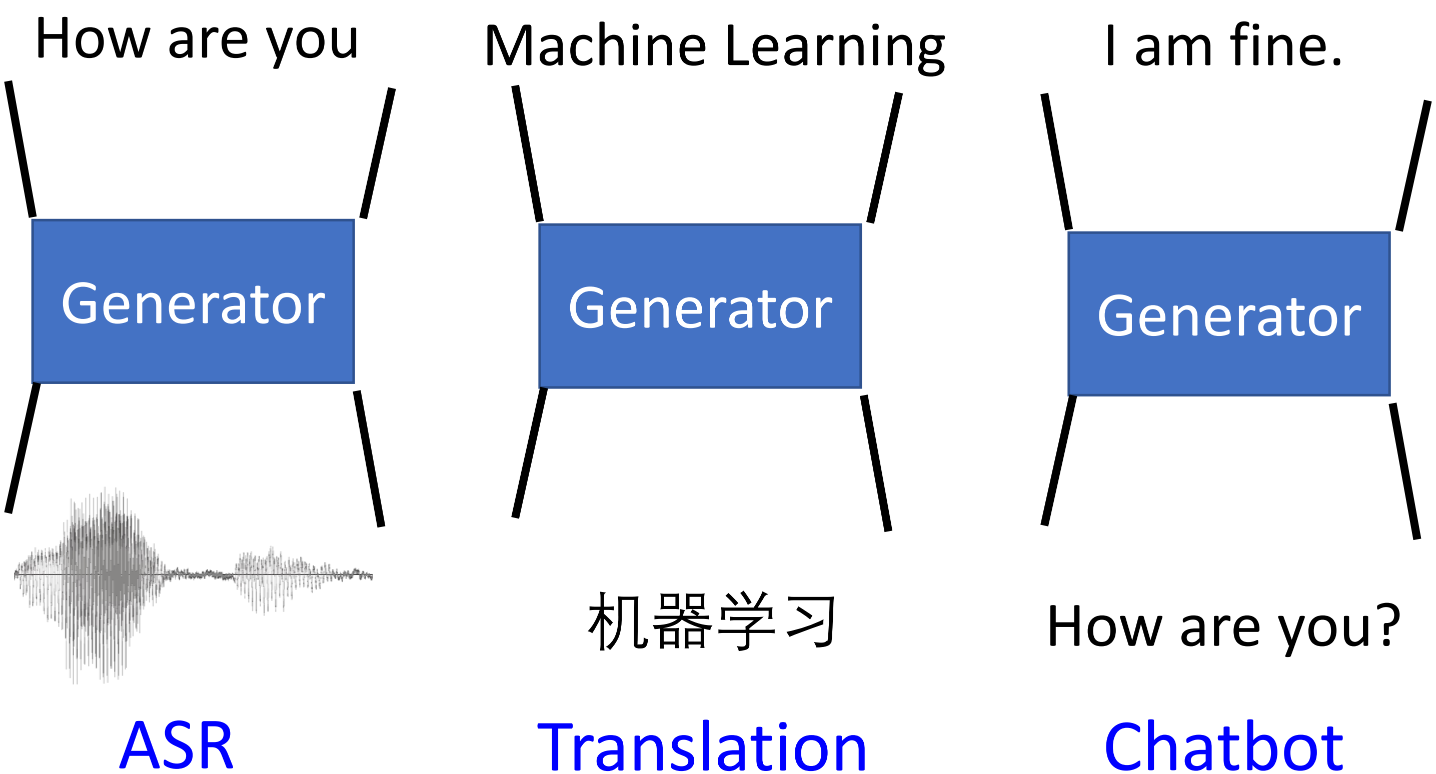
Conditional Sequence Generation

- RL (human feedback)
- GAN (discriminator feedback)

Unsupervised Conditional Sequence Generation

- Text Style Transfer
- Unsupervised Abstractive Summarization

Conditional Sequence Generation



The generator is a typical seq2seq model.

With GAN, you can train seq2seq model in another way.

Review: Sequence-to-sequence

- Chat-bot as example

Output:	Not bad	I'm John.
Human	better	
Training Criterion		better

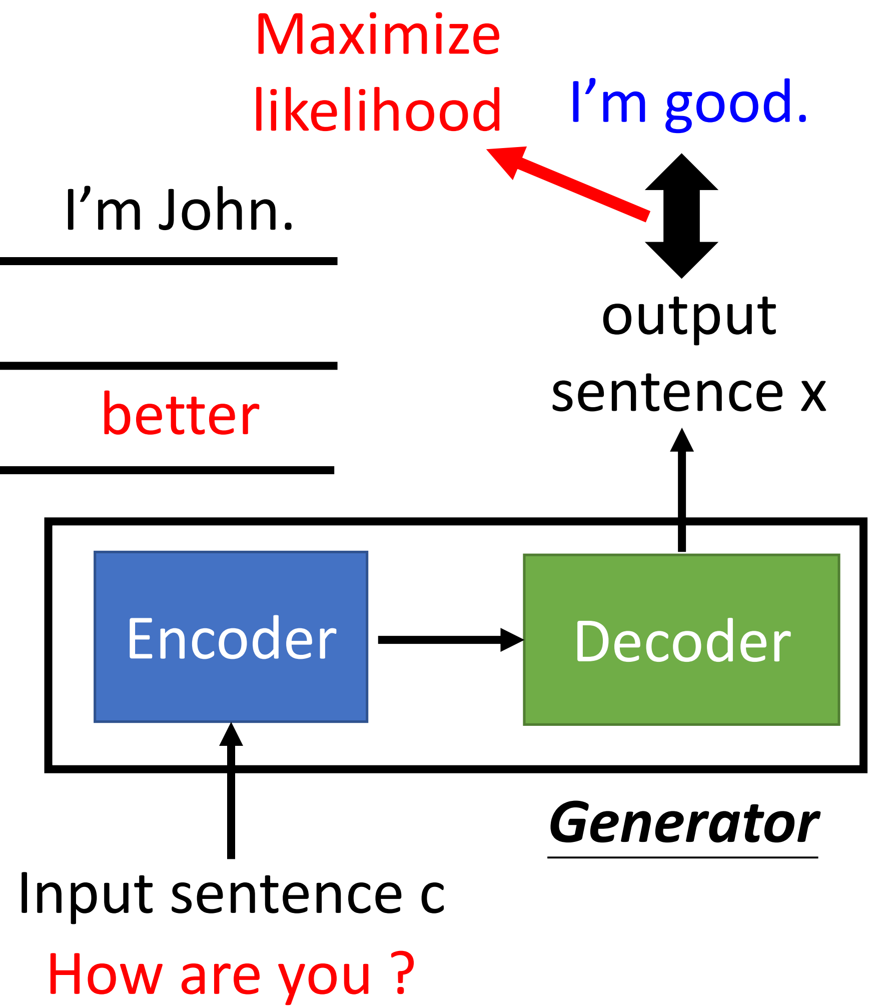
Training data:

⋮

A: How are you ?

B: I'm good.

⋮



Outline of Part III

Improving Supervised Seq-to-seq Model

- RL (human feedback)
- GAN (discriminator feedback)

Unsupervised Seq-to-seq Model

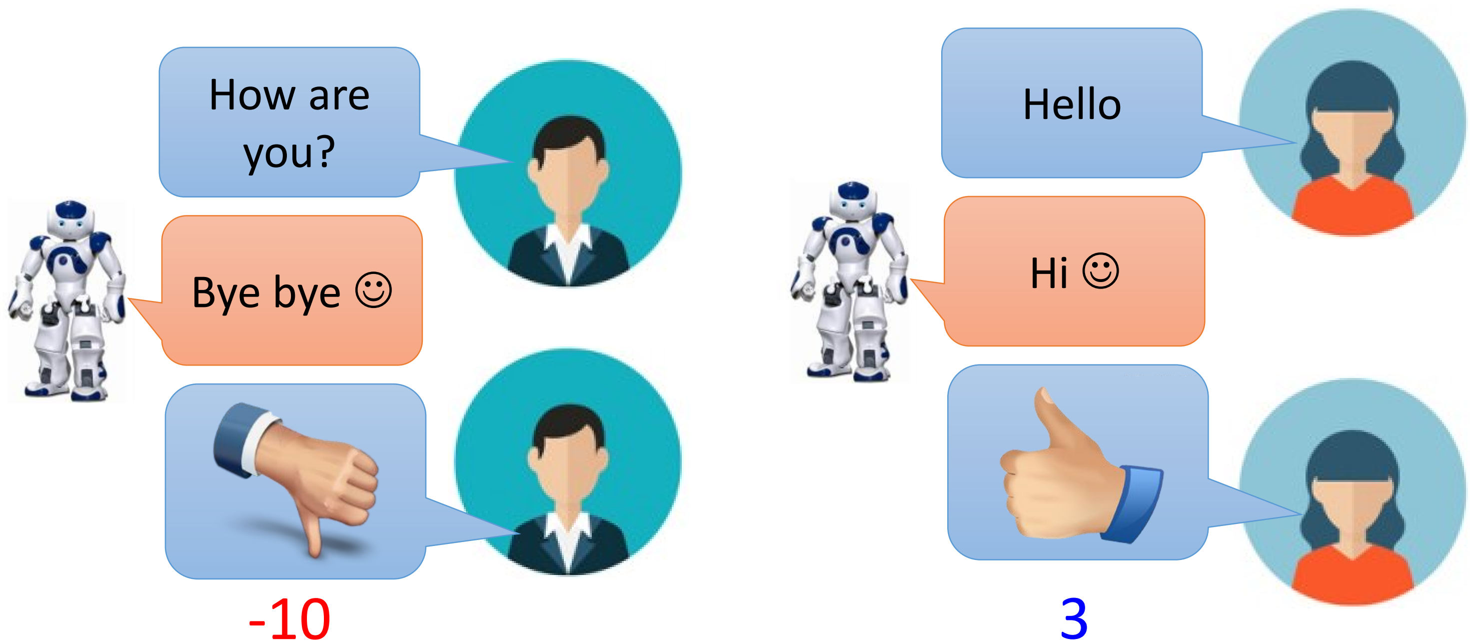
- Text Style Transfer
- Unsupervised Abstractive Summarization

https://image.freepik.com/free-vector/variety-of-human-avatars_23-2147506285.jpg

http://www.freepik.com/free-vector/variety-of-human-avatars_766615.htm

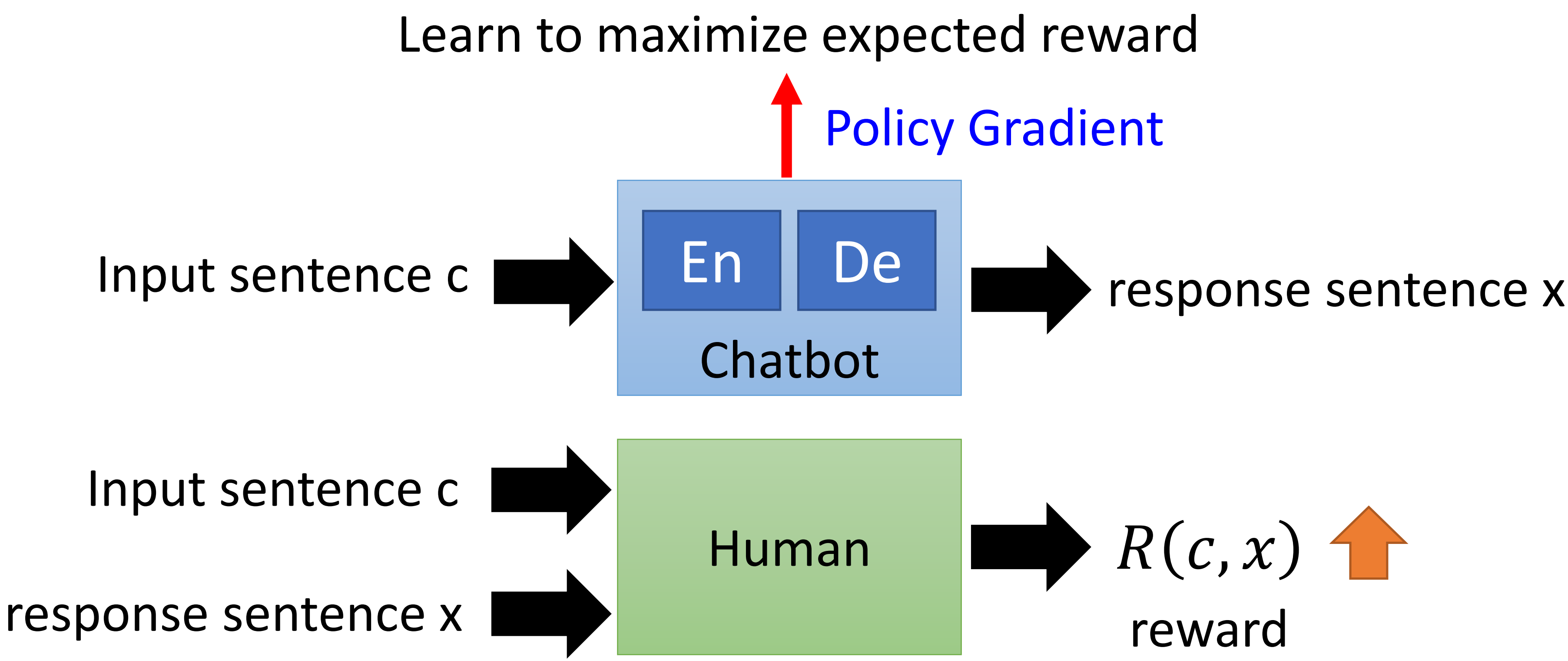
Introduction

- Machine obtains feedback from user



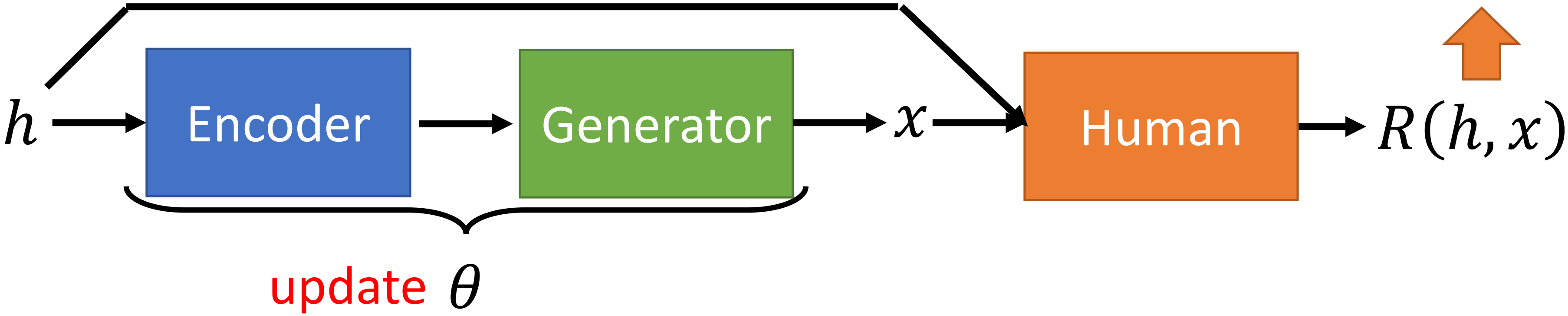
- Chat-bot learns to maximize the ***expected reward***

Maximizing Expected Reward



[Li, et al., EMNLP, 2016]

Maximizing Expected Reward



$$\theta^* = arg\max_{\theta} \bar{R}_\theta$$

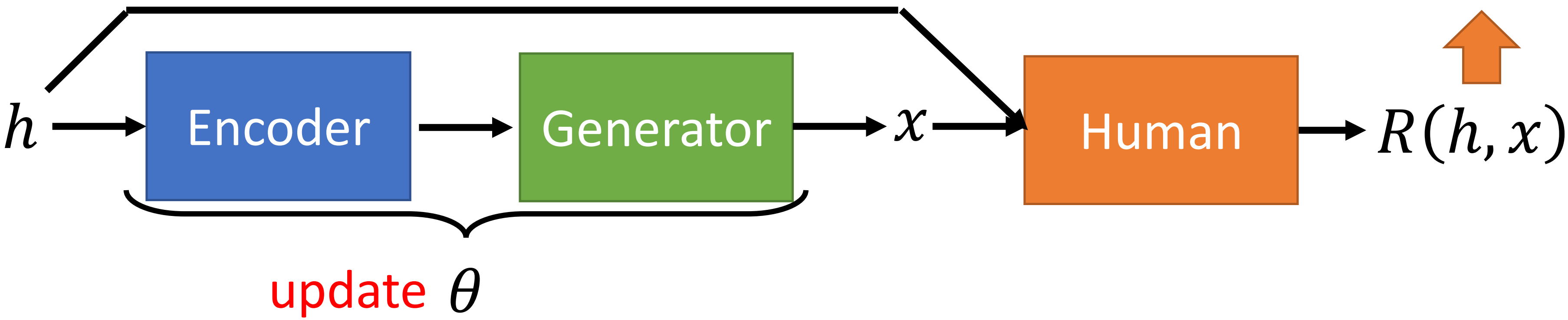
Maximizing expected reward

$$\bar{R}_\theta = \sum_h \underline{P(h)} \sum_x R(h,x) \, \underline{P_\theta(x|h)}$$

Randomness in generator

Probability that the input/history is h

Maximizing Expected Reward



$\theta^* = arg\max_{\theta} \bar{R}_\theta$ Maximizing expected reward

$$\bar{R}_\theta = \sum_h P(h) \sum_x R(h,x)\, P_\theta(x|h) = E_{h\sim P(h)}\left[E_{x\sim P_\theta(x|h)}[R(h,x)]\right]$$

$$= E_{h\sim P(h),x\sim P_\theta(x|h)}[R(h,x)] \approx \frac{1}{N}\sum_{i=1}^{N} R\left(h^i, x^i\right)$$

Sample: $(h^1, x^1), (h^2, x^2), \cdots, (h^N, x^N)$

Where is θ?

Policy Gradient

$$\boxed{\frac{d log(f(x))}{dx} = \frac{1}{f(x)} \frac{df(x)}{dx}}$$

$$\bar{R}_\theta = \sum_h P(h) \sum_x R(h,x)\, P_\theta(x|h) \approx \frac{1}{N} \sum_{i=1}^{N} R(h^i, x^i)$$

$$\nabla \bar{R}_\theta = \sum_h P(h) \sum_x R(h,x) \nabla P_\theta(x|h) \approx \frac{1}{N} \sum_{i=1}^{N} R(h^i, x^i) \nabla log P_\theta(x|h)$$

Sampling

$$= \sum_h P(h) \sum_x R(h,x) P_\theta(x|h) \boxed{\frac{\nabla P_\theta(x|h)}{P_\theta(x|h)}}$$

$$= \sum_h P(h) \sum_x R(h,x) P_\theta(x|h) \boxed{\nabla log P_\theta(x|h)}$$

$$= E_{h \sim P(h), x \sim P_\theta(x|h)} [R(h,x) \nabla log P_\theta(x|h)]$$

Policy Gradient

- Gradient Ascent

$$\theta^{new} \leftarrow \theta^{old} + \eta \nabla \bar{R}_{\theta^{old}}$$

$$\nabla \bar{R}_{\theta} \approx \frac{1}{N} \sum_{i=1}^{N} R\left(h^i, x^i\right) \nabla log P_{\theta}\left(x^i | h^i\right)$$

$R\left(h^i, x^i\right)$ is positive

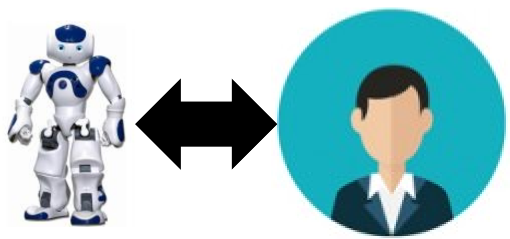
After updating θ, $P_{\theta}\left(x^i | h^i\right)$ will increase

$R\left(h^i, x^i\right)$ is negative

After updating θ, $P_{\theta}\left(x^i | h^i\right)$ will decrease

Policy Gradient - Implemenation

θ^t



(c^1, x^1) $\quad R(c^1, x^1)$

(c^2, x^2) $\quad R(c^2, x^2)$

$\vdots \qquad \vdots$

(c^N, x^N) $\quad R(c^N, x^N)$

$$\theta^{t+1} \leftarrow \theta^t + \eta \boxed{\nabla \bar{R}_{\theta^t}}$$

$$\frac{1}{N}\sum_{i=1}^{N} R(c^i, x^i) \nabla log P_{\theta^t}(x^i|c^i)$$

$R(c^i, x^i)$ is positive

Updating θ to increase $P_\theta(x^i|c^i)$

$R(c^i, x^i)$ is negative

Updating θ to decrease $P_\theta(x^i|c^i)$

Comparison

	Maximum Likelihood	Reinforcement Learning
Objective Function	$\frac{1}{N}\sum_{i=1}^{N} logP_\theta(\hat{x}^i \mid c^i)$	$\frac{1}{N}\sum_{i=1}^{N} R(c^i, x^i) logP_\theta(x^i \mid c^i)$
Gradient	$\frac{1}{N}\sum_{i=1}^{N} \nabla logP_\theta(\hat{x}^i \mid c^i)$	$\frac{1}{N}\sum_{i=1}^{N} R(c^i, x^i) \nabla logP_\theta(x^i \mid c^i)$
Training Data	$\{(c^1, \hat{x}^1), \dots, (c^N, \hat{x}^N)\}$ $R(c^i, \hat{x}^i) = 1$	$\{(c^1, x^1), \dots, (c^N, x^N)\}$ obtained from interaction weighted by $R(c^i, x^i)$

Alpha GO style training !

- Let two agents talk to each other

How old are you?

See you.

See you.

See you.

How old are you?

I am 16.

I though you were 12.

What make you think so?

Using a pre-defined evaluation function to compute R(h,x)

Outline of Part III

Improving Supervised Seq-to-seq Model

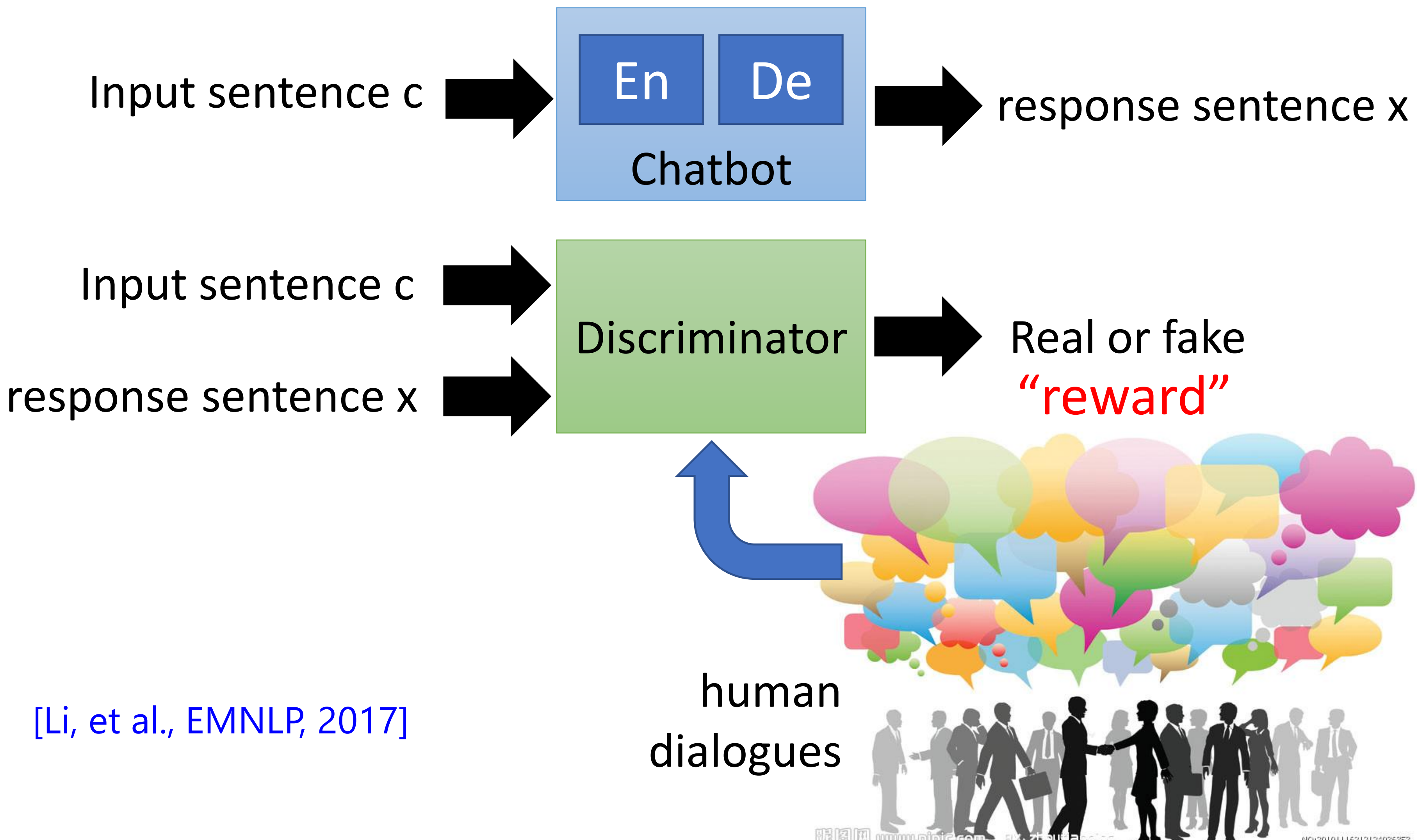
- RL (human feedback)
- GAN (discriminator feedback)

Unsupervised Seq-to-seq Model

- Text Style Transfer
- Unsupervised Abstractive Summarization

http://www.nipic.com/show/3/83/3936650kd7476069.html

Conditional GAN



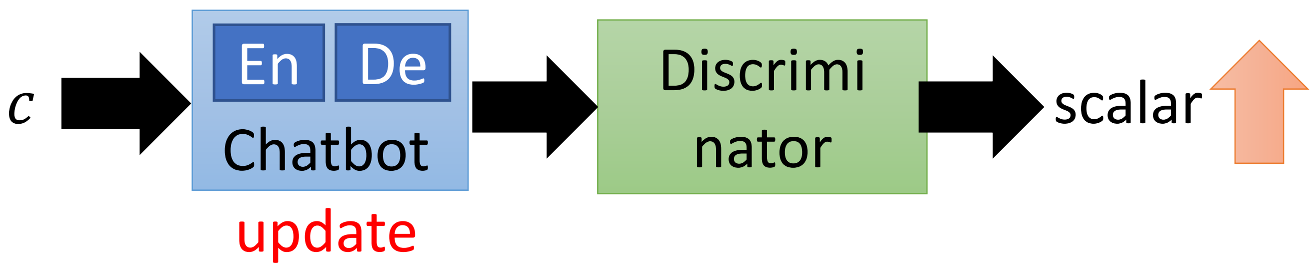
[Li, et al., EMNLP, 2017]

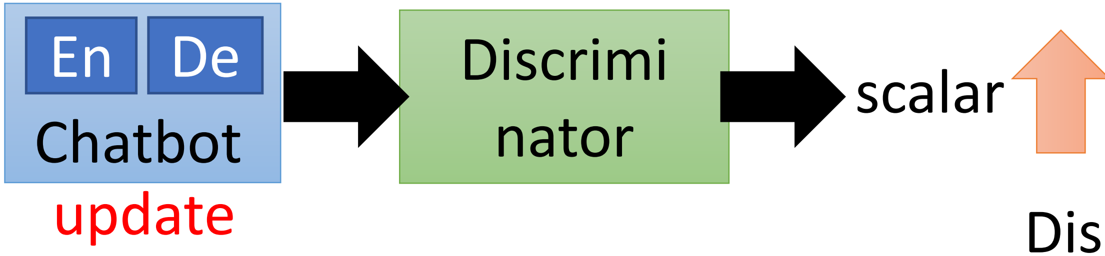
Algorithm

Training data:

Pairs of conditional input c and response x

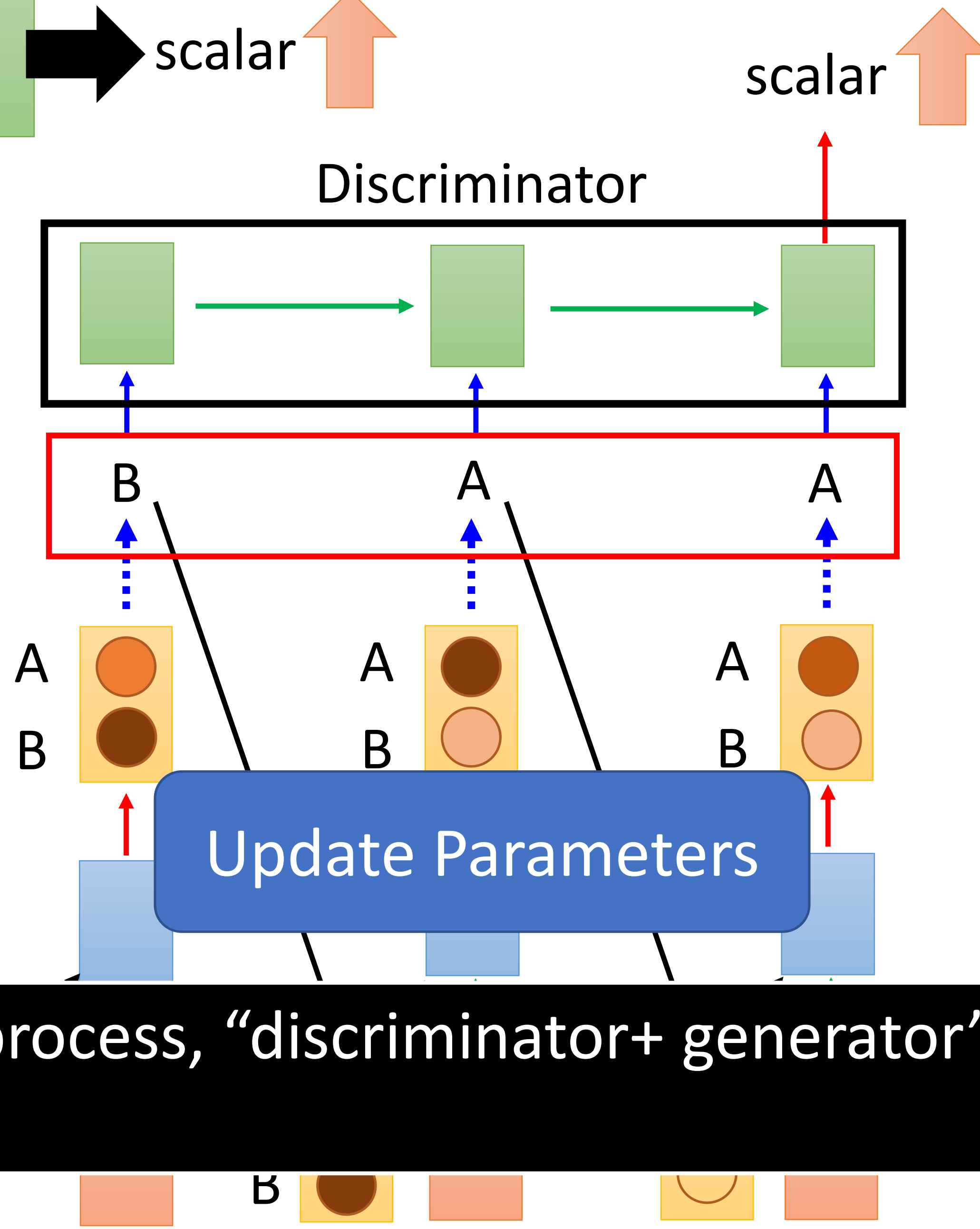
- Initialize generator G (chatbot) and discriminator D
- In each iteration:
  - Sample input c and response x from training set
  - Sample input c' from training set, and generate response $\tilde{x}$ by G(c')
  - Update D to increase $D(c, x)$ and decrease $D(c', \tilde{x})$
  - Update generator G (chatbot) such that





Can we use gradient ascent?

NO!



Due to the sampling process, “discriminator+ generator” is not differentiable

Three Categories of Solutions

Gumbel-softmax

- [Matt J. Kusner, et al, arXiv, 2016]

Continuous Input for Discriminator

- [Sai Rajeswar, et al., arXiv, 2017][Ofir Press, et al., ICML workshop, 2017][Zhen Xu, et al., EMNLP, 2017][Alex Lamb, et al., NIPS, 2016][Yizhe Zhang, et al., ICML, 2017]

“Reinforcement Learning”

- [Yu, et al., AAAI, 2017][Li, et al., EMNLP, 2017][Tong Che, et al, arXiv, 2017][Jiaxian Guo, et al., AAAI, 2018][Kevin Lin, et al, NIPS, 2017][William Fedus, et al., ICLR, 2018]

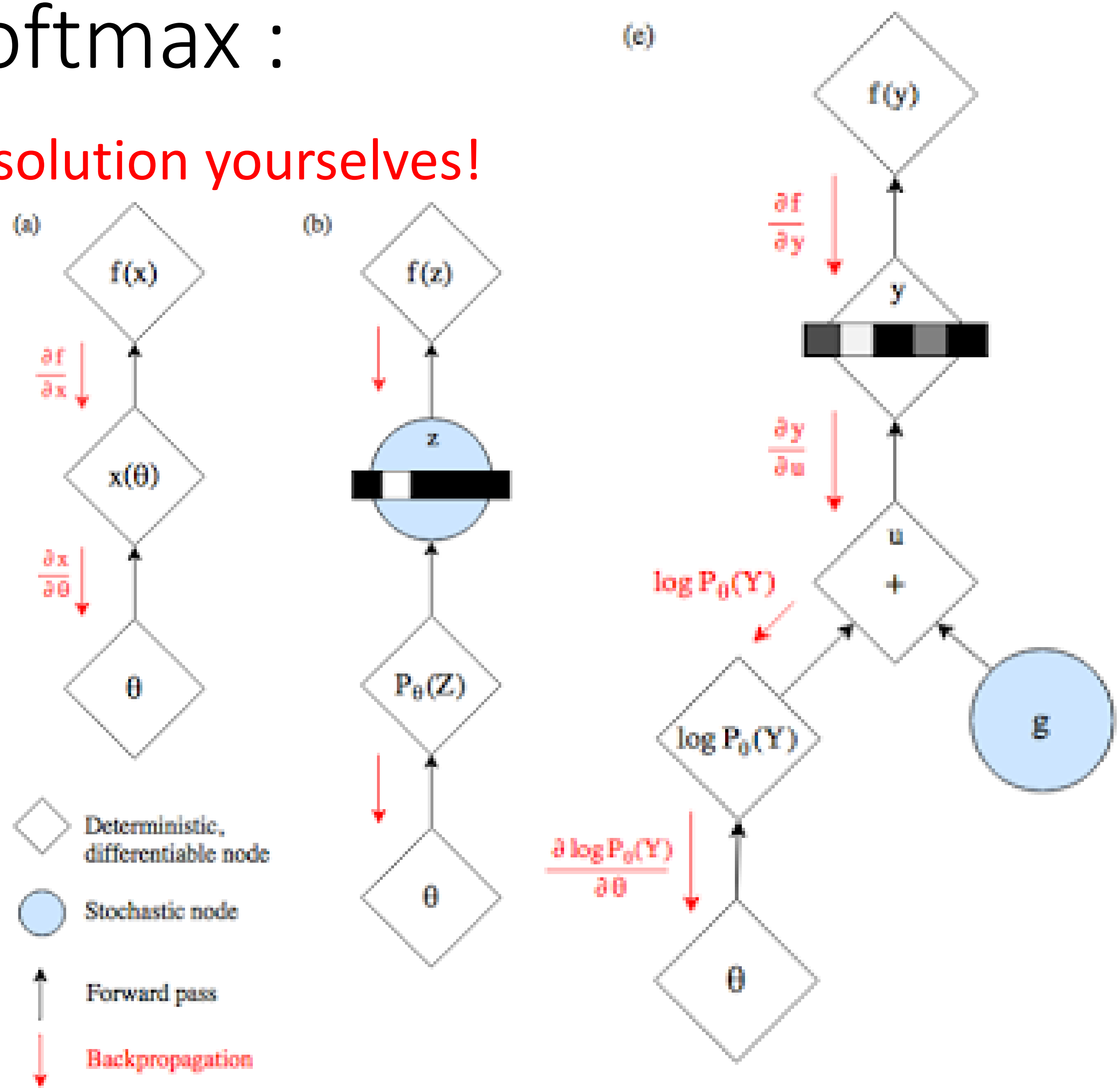
Gumbel-softmax :

How? Find the solution yourselves!

https://gabrielhuang.gitbooks.io/machine-learning/reparametrization-trick.html

https://casmls.github.io/general/2017/02/01/GumbelSoftmax.html

http://blog.evjang.com/2016/11/tutorial-categorical-variational.html



Three Categories of Solutions

Gumbel-softmax

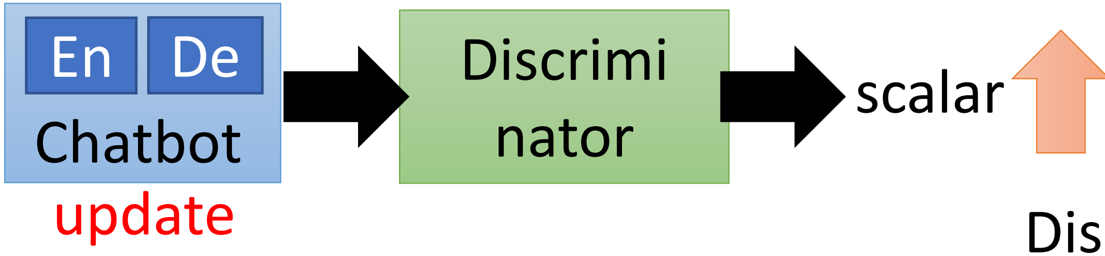
- [Matt J. Kusner, et al, arXiv, 2016]

Continuous Input for Discriminator

- [Sai Rajeswar, et al., arXiv, 2017][Ofir Press, et al., ICML workshop, 2017][Zhen Xu, et al., EMNLP, 2017][Alex Lamb, et al., NIPS, 2016][Yizhe Zhang, et al., ICML, 2017]

“Reinforcement Learning”

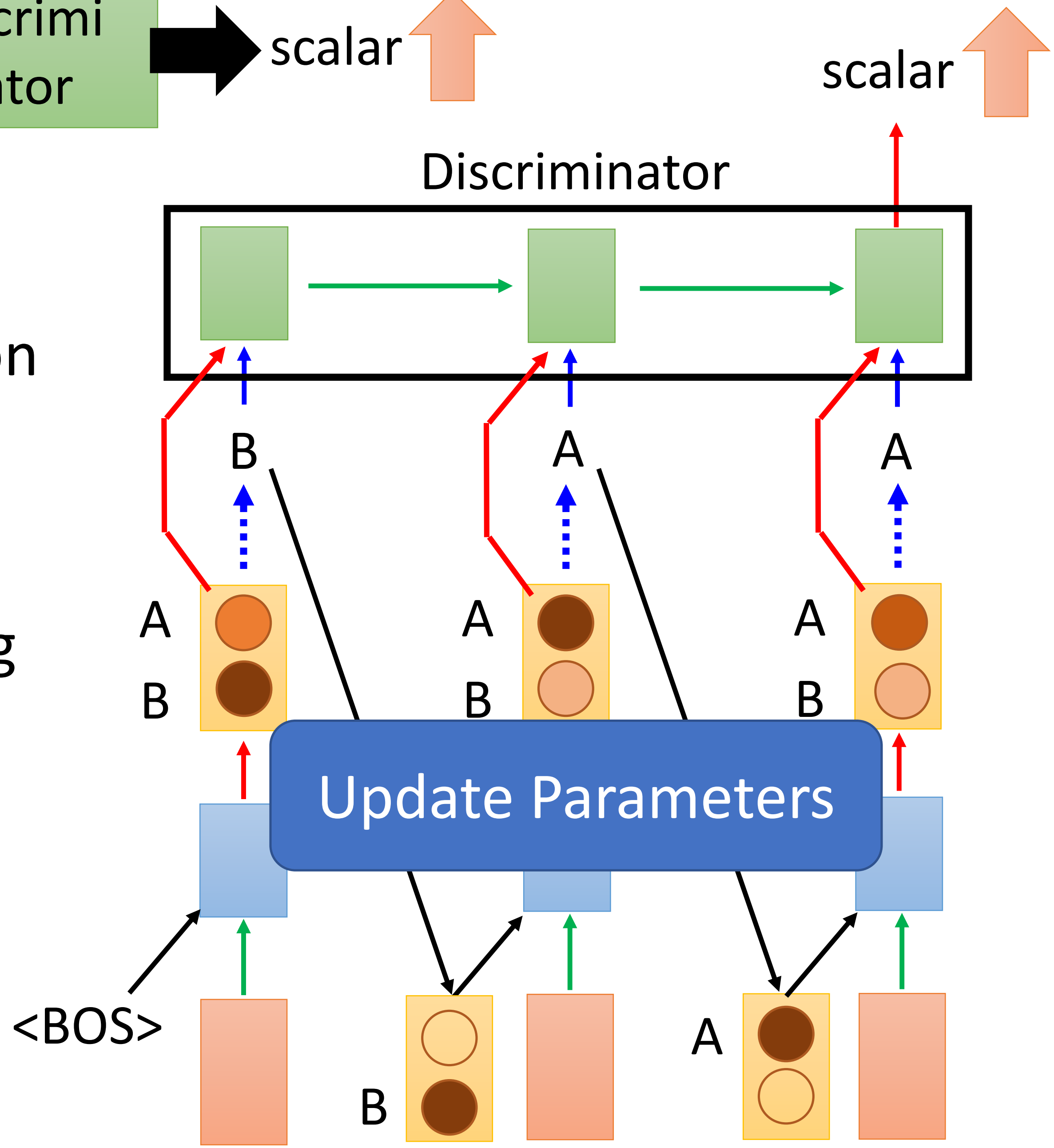
- [Yu, et al., AAAI, 2017][Li, et al., EMNLP, 2017][Tong Che, et al, arXiv, 2017][Jiaxian Guo, et al., AAAI, 2018][Kevin Lin, et al, NIPS, 2017][William Fedus, et al., ICLR, 2018]



Use the distribution as the input of discriminator

Avoid the sampling process

We can do backpropagation now.



Three Categories of Solutions

Gumbel-softmax

- [Matt J. Kusner, et al, arXiv, 2016]

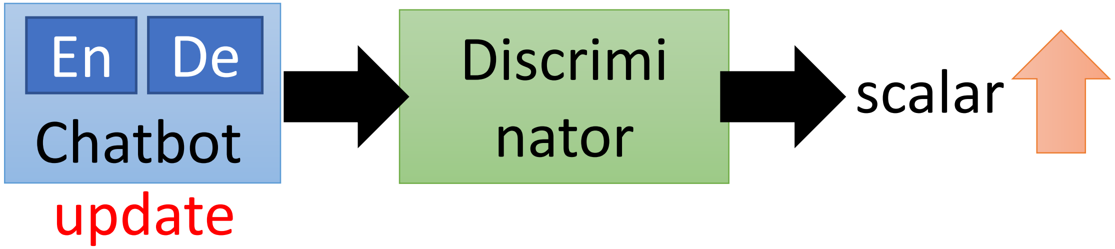
Continuous Input for Discriminator

- [Sai Rajeswar, et al., arXiv, 2017][Ofir Press, et al., ICML workshop, 2017][Zhen Xu, et al., EMNLP, 2017][Alex Lamb, et al., NIPS, 2016][Yizhe Zhang, et al., ICML, 2017]

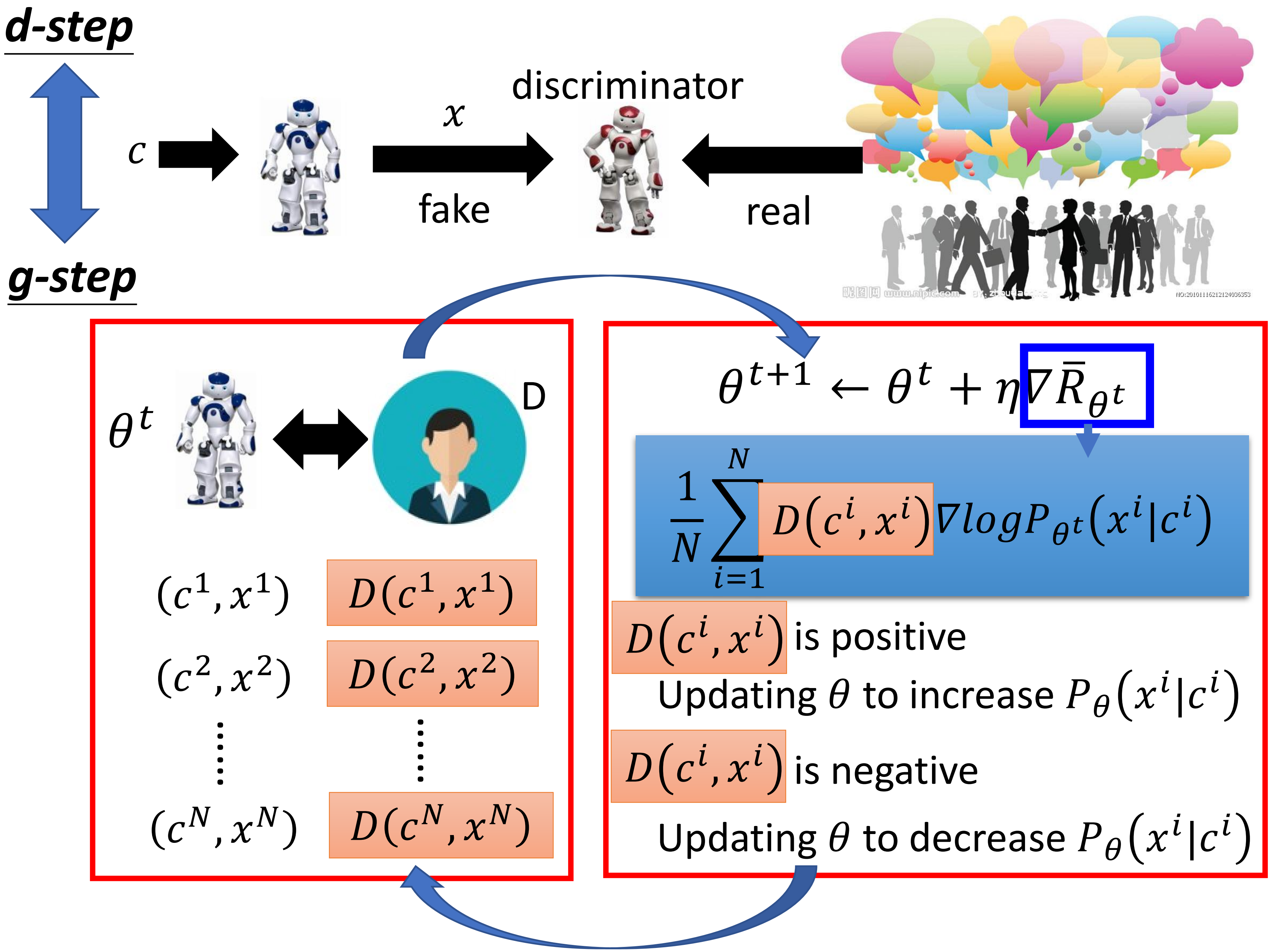
“Reinforcement Learning”

- [Yu, et al., AAAI, 2017][Li, et al., EMNLP, 2017][Tong Che, et al, arXiv, 2017][Jiaxian Guo, et al., AAAI, 2018][Kevin Lin, et al, NIPS, 2017][William Fedus, et al., ICLR, 2018]

Reinforcement Learning?



- Consider the output of discriminator as reward
  - Update generator to increase discriminator = to get maximum reward
  - Using the formulation of policy gradient, replace reward $R(c,x)$ with discriminator output $\mathrm{D}(c,x)$
- Different from typical RL
  - The discriminator would update


d-step
g-step
c
x
discriminator
fake
real
θ^t
D
(c^1, x^1)
$D(c^1, x^1)$
(c^2, x^2)
$D(c^2, x^2)$
(c^N, x^N)
$D(c^N, x^N)$
$\theta^{t+1} \leftarrow \theta^t + \eta \boxed{\nabla \bar{R}_{\theta^t}}$
$\frac{1}{N} \sum_{i=1}^{N} D(c^i, x^i) \nabla log P_{\theta^t}(x^i | c^i)$
$D(c^i, x^i)$ is positive
Updating θ to increase $P_\theta(x^i | c^i)$
$D(c^i, x^i)$ is negative
Updating θ to decrease $P_\theta(x^i | c^i)$

Reward for Every Generation Step

$$\nabla \bar{R}_\theta \approx \frac{1}{N}\sum_{i=1}^{N} D\left(c^i, x^i\right) \nabla logP_\theta\left(x^i | c^i\right)$$

c^i = "What is your name?"

x^i = "I don't know"

$D\left(c^i, x^i\right)$ is negative

Update θ to decrease $\log P_\theta\left(x^i | c^i\right)$

$$logP_\theta\left(x^i | c^i\right) = logP\left(x_1^i | c^i\right) + logP\left(x_2^i | c^i, x_1^i\right) + logP\left(x_3^i | c^i, x_{1:2}^i\right)$$

$P\left("I" | c^i\right)$?

c^i = "What is your name?"

x^i = "I am John"

$D\left(c^i, x^i\right)$ is positive

Update θ to increase $\log P_\theta\left(x^i | c^i\right)$

$$logP_\theta\left(x^i | c^i\right) = logP\left(x_1^i | c^i\right) + logP\left(x_2^i | c^i, x_1^i\right) + logP\left(x_3^i | c^i, x_{1:2}^i\right)$$

$P\left("I" | c^i\right)$

Reward for Every Generation Step

h^i = "What is your name?" x^i = "I don't know"

$$logP_\theta\left(x^i|h^i\right) = logP\left(x_1^i|c^i\right) + logP\left(x_2^i|c^i, x_1^i\right) + logP\left(x_3^i|c^i, x_{1:2}^i\right)$$

$$P\left("I"|c^i\right) \quad P\left("don't"|c^i, "I"\right) \quad P\left("know"|c^i, "I\ don't"\right)$$

$$\nabla\bar{R}_\theta \approx \frac{1}{N}\sum_{i=1}^{N} D\left(c^i, x^i\right)\nabla logP_\theta\left(x^i|c^i\right)$$

$$\nabla\bar{R}_\theta \approx \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(Q\left(c^i, x_{1:t}^i\right) - b\right)\nabla logP_\theta\left(x_t^i|c^i, x_{1:t-1}^i\right)$$

Method 1. Monte Carlo (MC) Search [Yu, et al., AAAI, 2017]

Method 2. Discriminator For Partially Decoded Sequences

[Li, et al., EMNLP, 2017]

Tips: RankGAN

Kevin Lin, Dianqi Li, Xiaodong He, Zhengyou Zhang, Ming-Ting Sun, "Adversarial Ranking for Language Generation", NIPS 2017

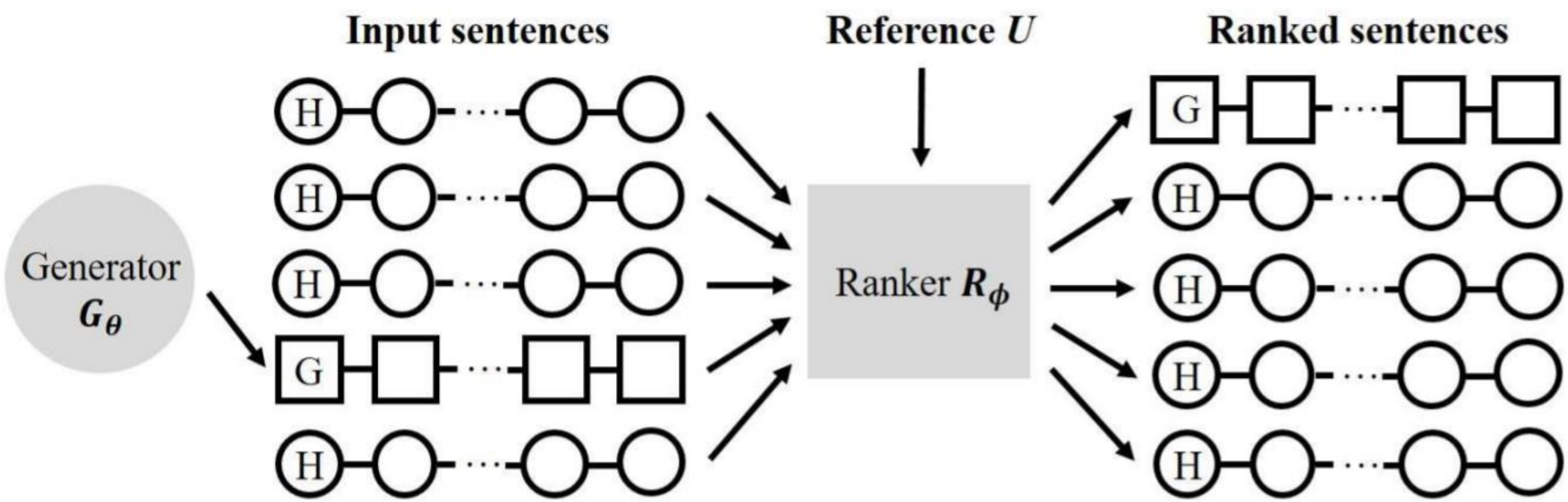


Image caption generation:

Method	BLEU-2	BLEU-3	BLEU-4
MLE	0.781	0.624	0.589
SeqGAN	0.815	0.636	0.587
RankGAN	**0.845**	**0.668**	**0.614**

Method	Human score
SeqGAN	3.44
RankGAN	4.61
Human-written	**6.42**

Experimental Results

Input	We've got to look for another route.
MLE	I'm sorry.
GAN	You're not going to be here for a while.

Input	You can save him by talking.
MLE	I don't know.
GAN	You know what's going on in there, you know what I mean?

- MLE frequently generates “I’m sorry”, “I don’t know”, etc. (corresponding to fuzzy images?)
- GAN generates longer and more complex responses (however, no strong evidence shows that they are better)

Find more comparison in the survey papers.

[Lu, et al., arXiv, 2018][Zhu, et al., arXiv, 2018]

More Applications

- Supervised machine translation [Wu, et al., arXiv 2017][Yang, et al., arXiv 2017]
- Supervised abstractive summarization [Liu, et al., AAAI 2018]
- Image/video caption generation [Rakshith Shetty, et al., ICCV 2017][Liang, et al., arXiv 2017]

If you are using seq2seq models,
consider to improve them by GAN.

Outline of Part III

Conditional Sequence Generation

- RL (human feedback)
- GAN (discriminator feedback)

Unsupervised Conditional Sequence Generation

- Text Style Transfer
- Unsupervised Abstractive Summarization

Text Style Transfer

Domain X

male

Domain Y

female

It is good.

It's a good day.

I love you.

positive sentences

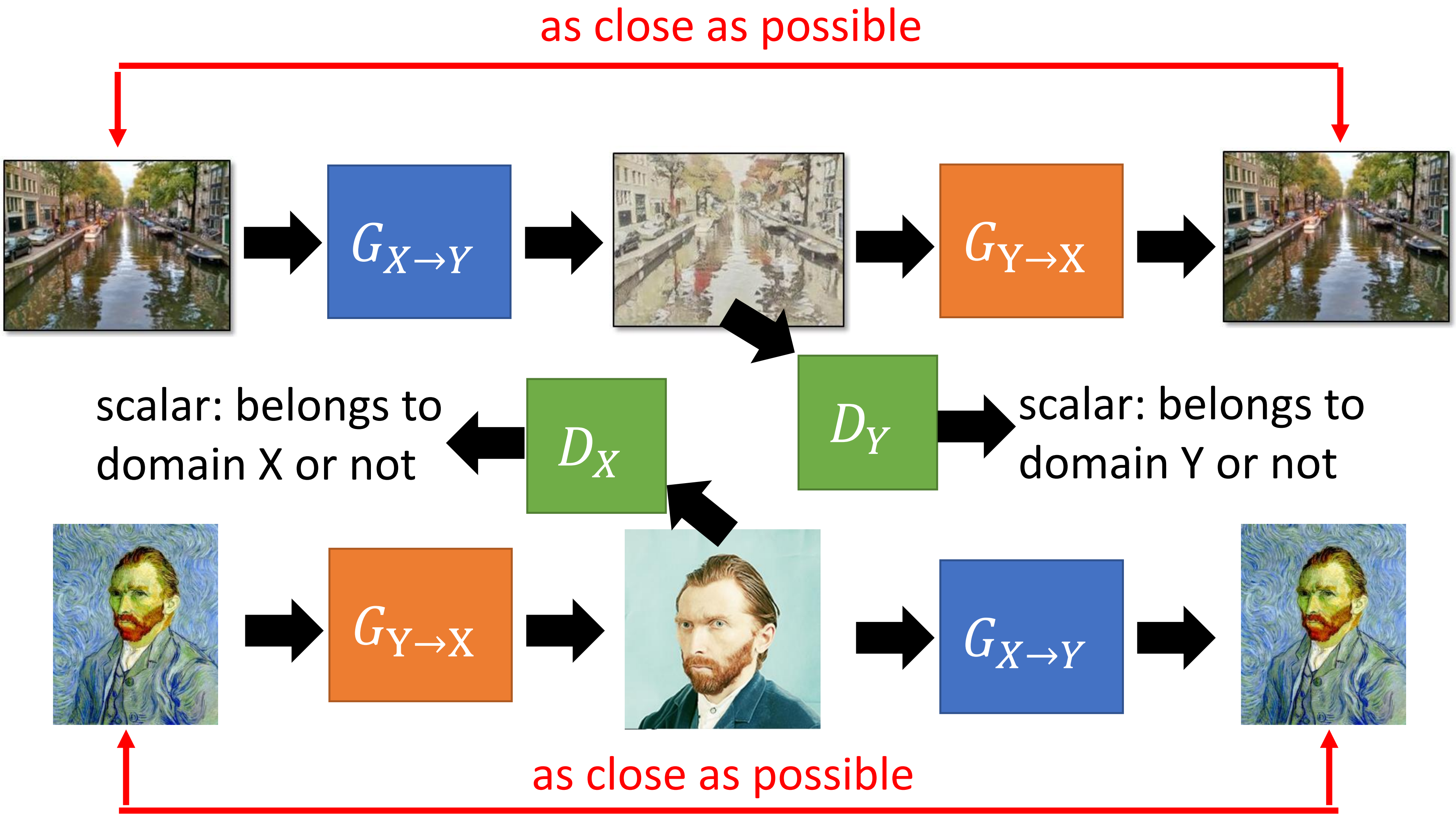
It is bad.

It's a bad day.

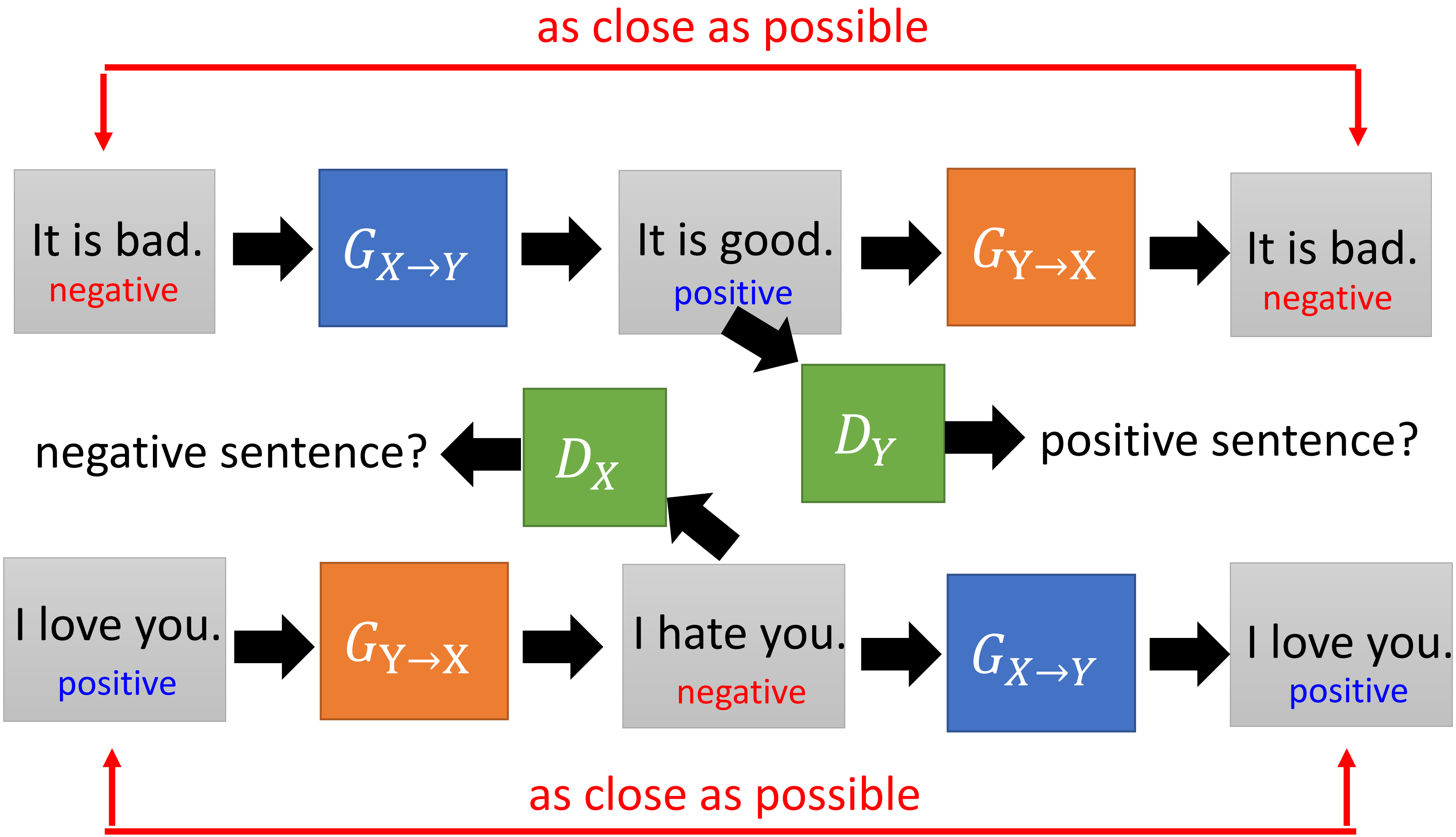
I don't love you.

negative sentences

Direct Transformation



Direct Transformation

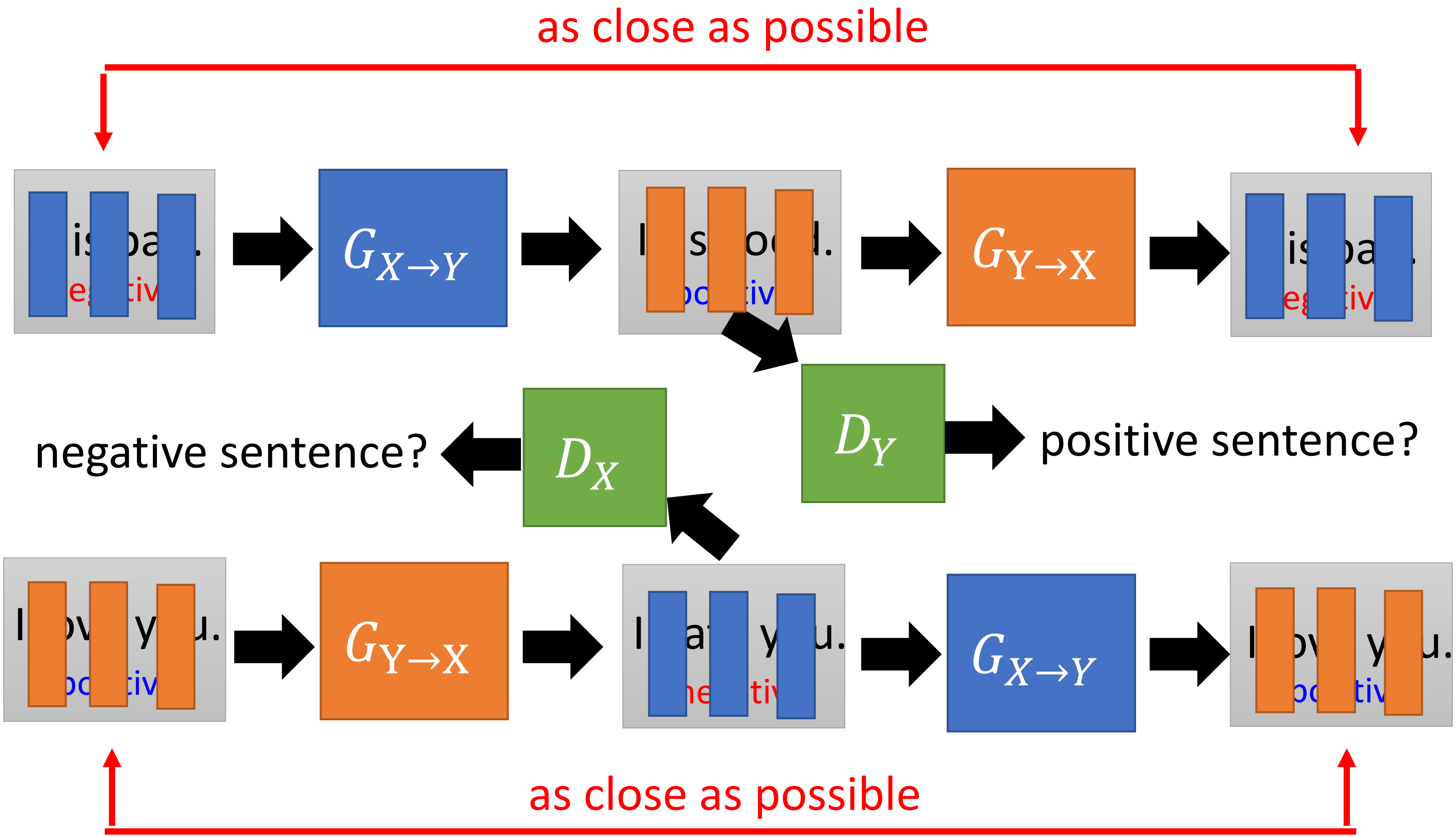


Direct Transformation

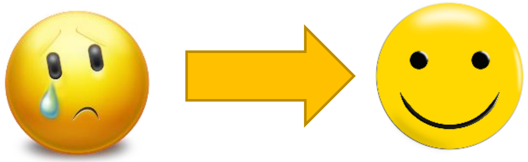
Discrete?

Word embedding

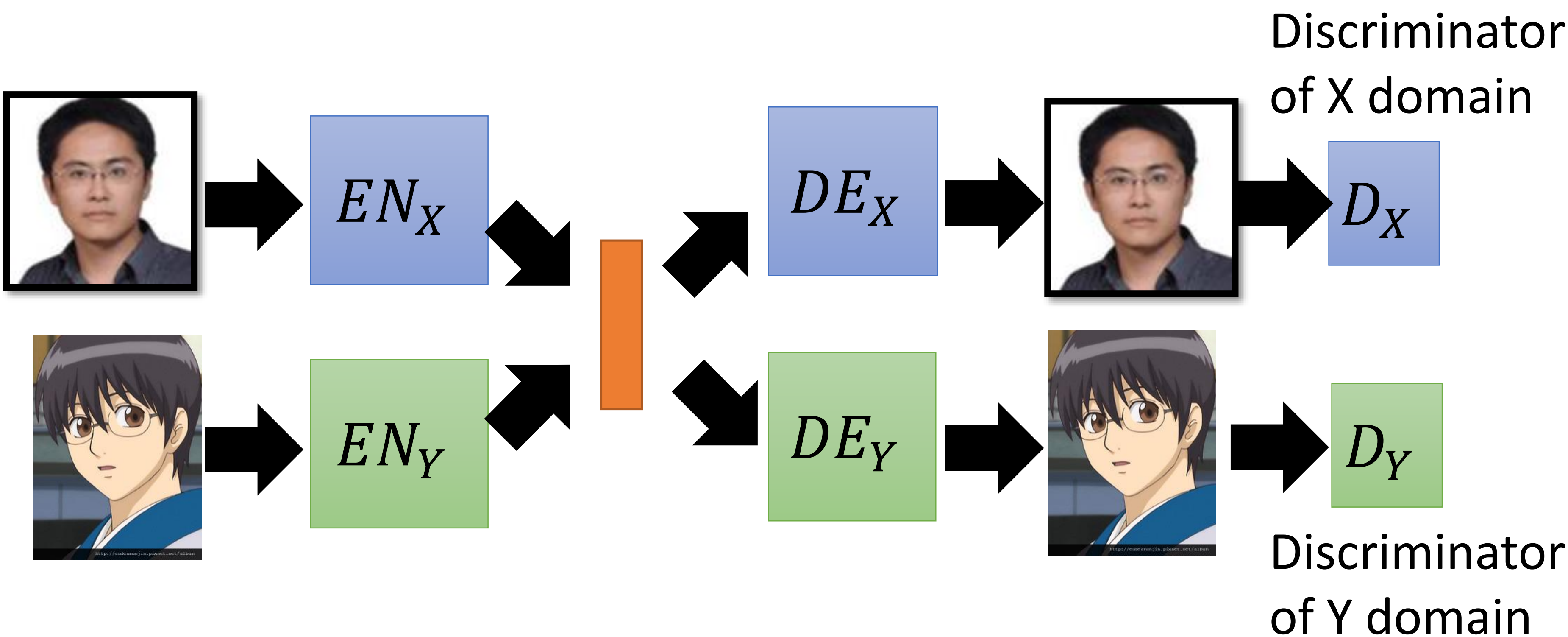
[Lee, et al., ICASSP, 2018]



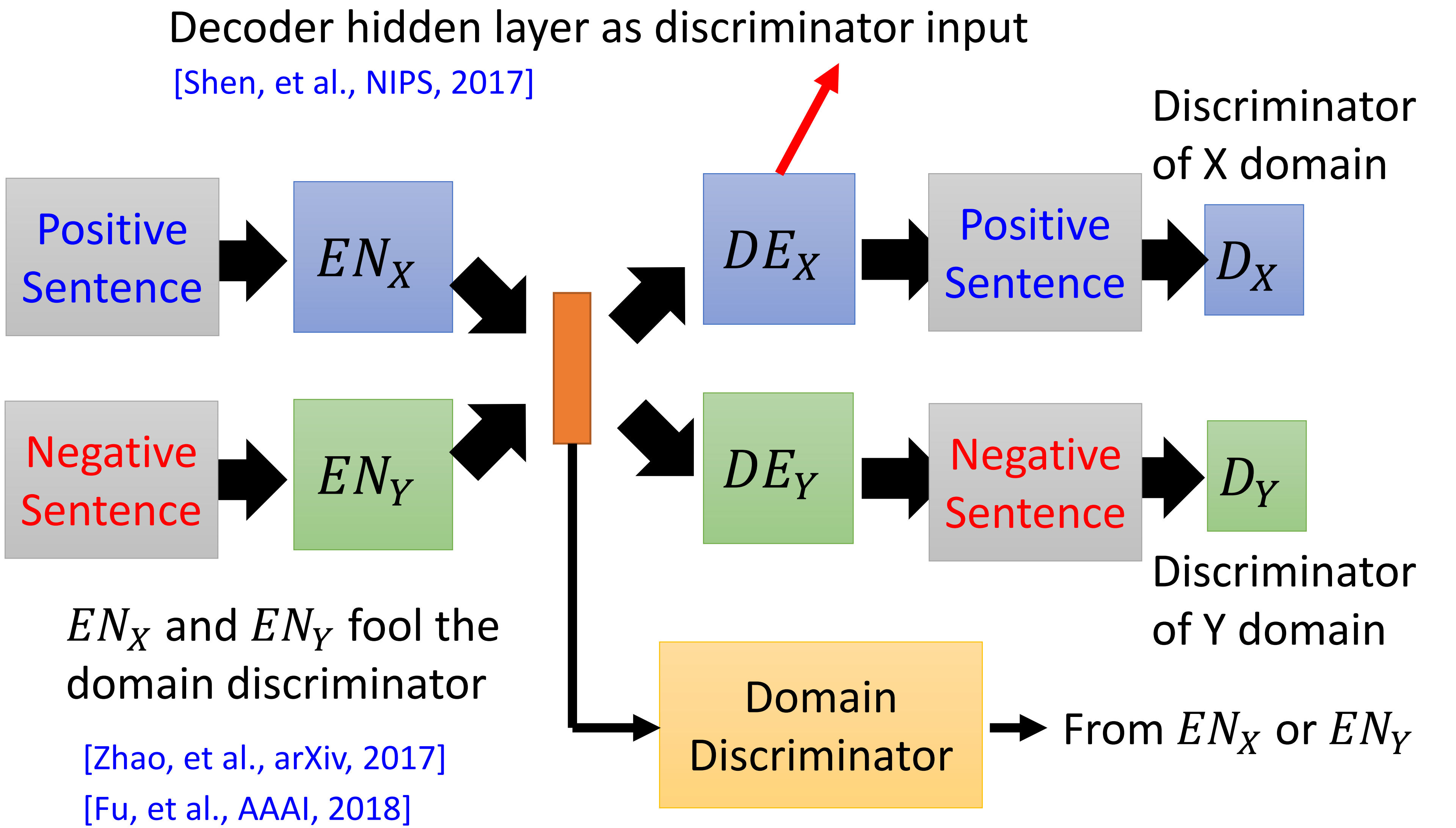
- Negative sentence to positive sentence:
  it's a crappy day → it's a great day
  i wish you could be here → you could be here
  it's not a good idea → it's good idea
  i miss you → i love you
  i don't love you → i love you
  i can't do that → i can do that
  i feel so sad → i happy
  it's a bad day → it's a good day
  it's a dummy day → it's a great day
  sorry for doing such a horrible thing → thanks for doing a great thing
  my doggy is sick → my doggy is my doggy
  my little doggy is sick → my little doggy is my little doggy



Projection to Common Space



Projection to Common Space



Outline of Part III

Improving Supervised Seq-to-seq Model

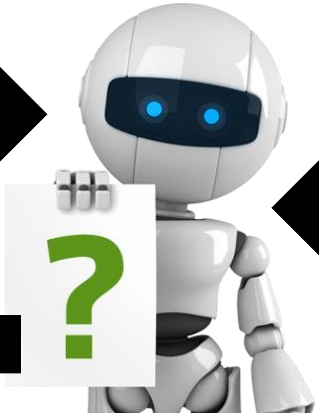
- RL (human feedback)
- GAN (discriminator feedback)

Unsupervised Seq-to-seq Model

- Text Style Transfer
- Unsupervised Abstractive Summarization

Abstractive Summarization

- Now machine can do **abstractive summary** by seq2seq (write summaries in its own words)



summary 1

summary 2

summary 3

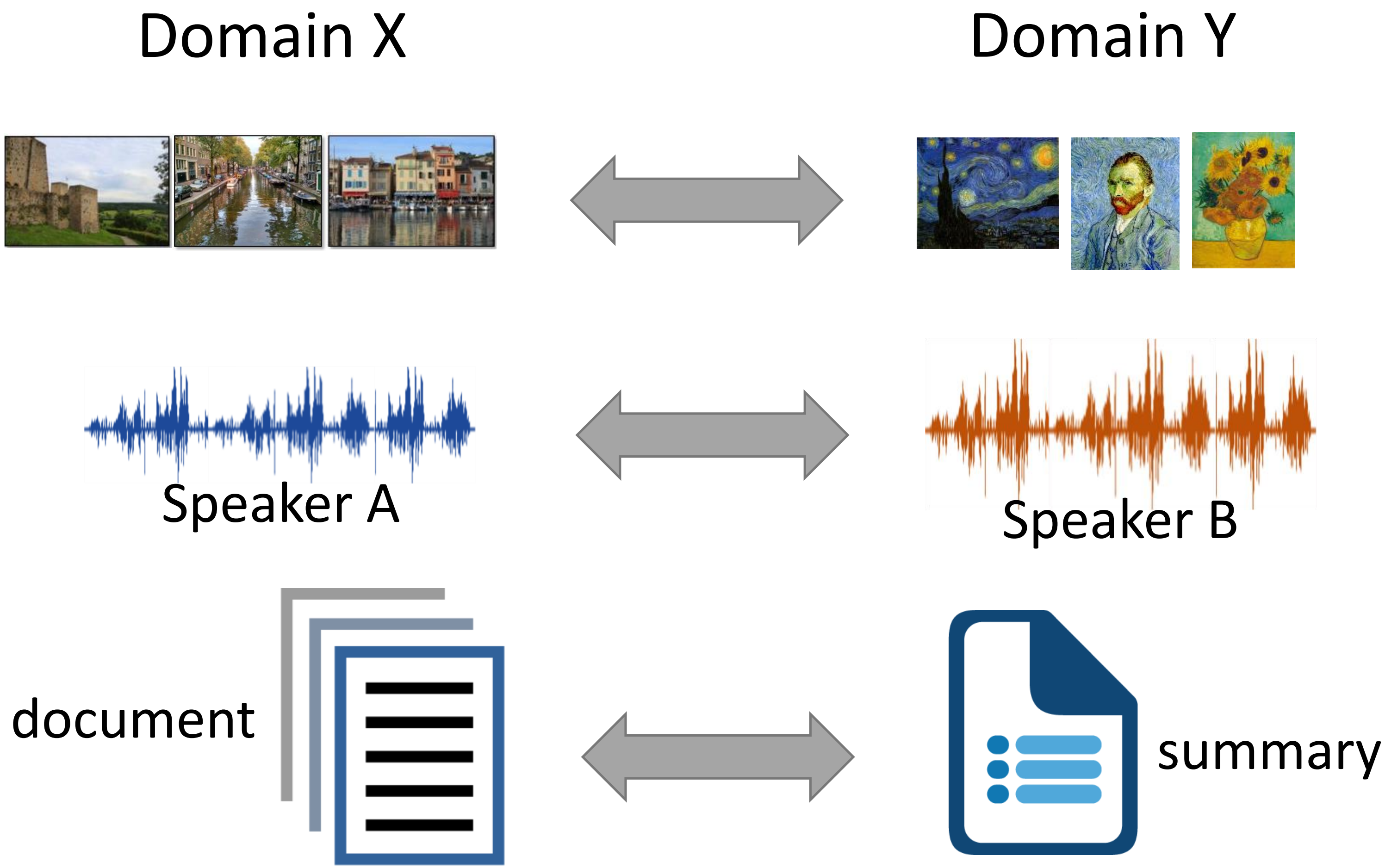
summary
(in its own words)

seq2seq

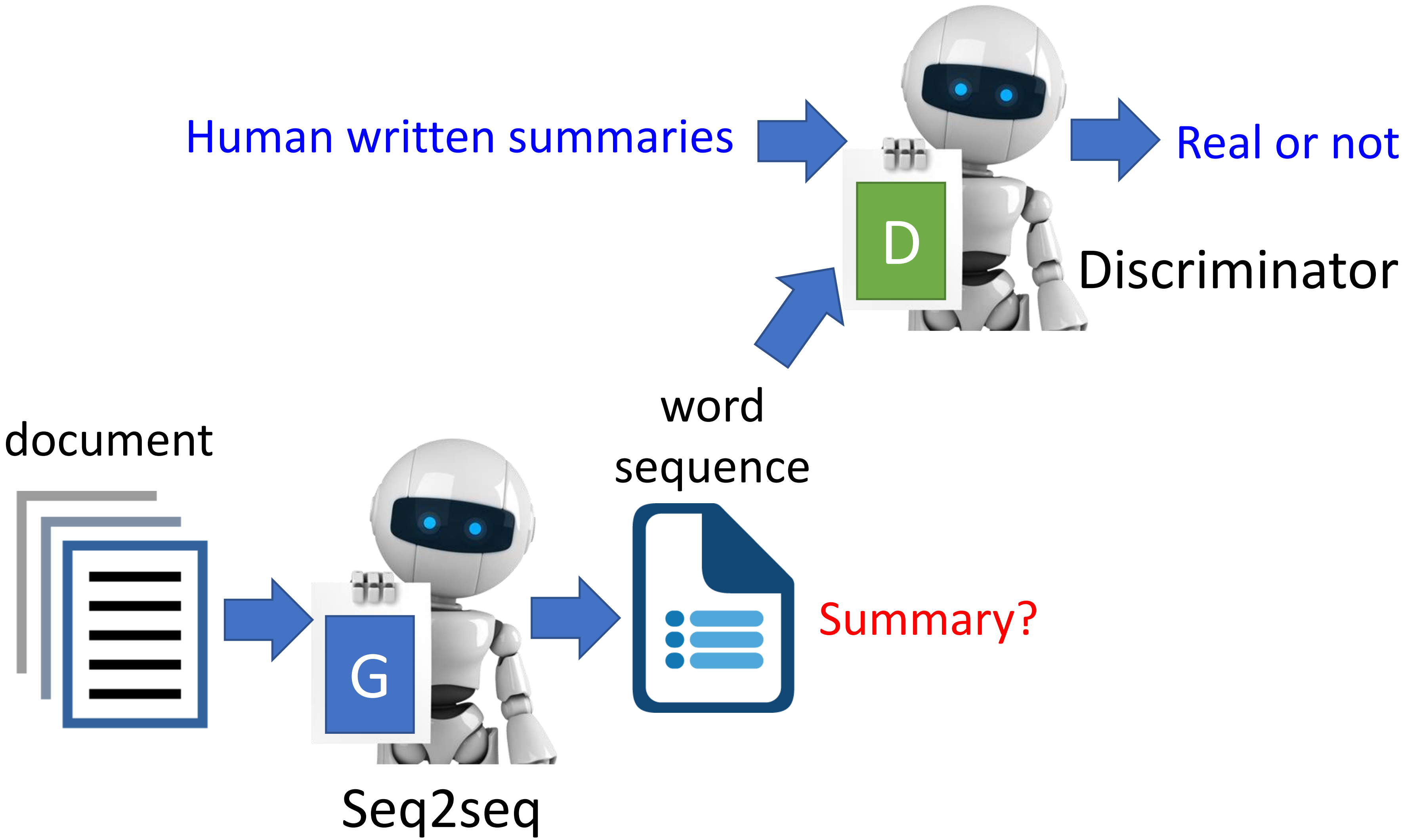
Training Data

Supervised: We need lots of labelled training data.

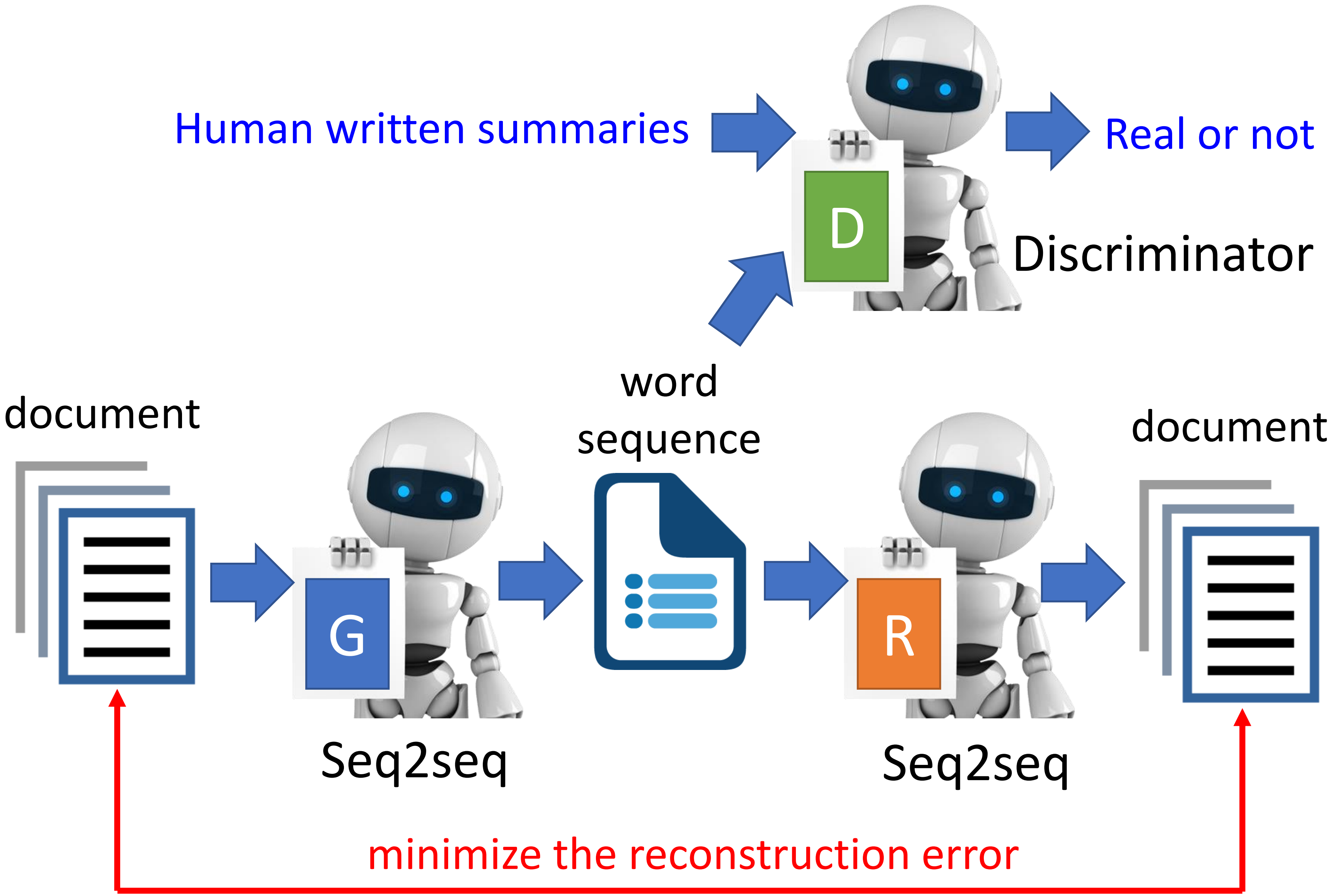
Unsupervised Conditional Generation



Unsupervised Abstractive Summarization



Unsupervised Abstractive Summarization

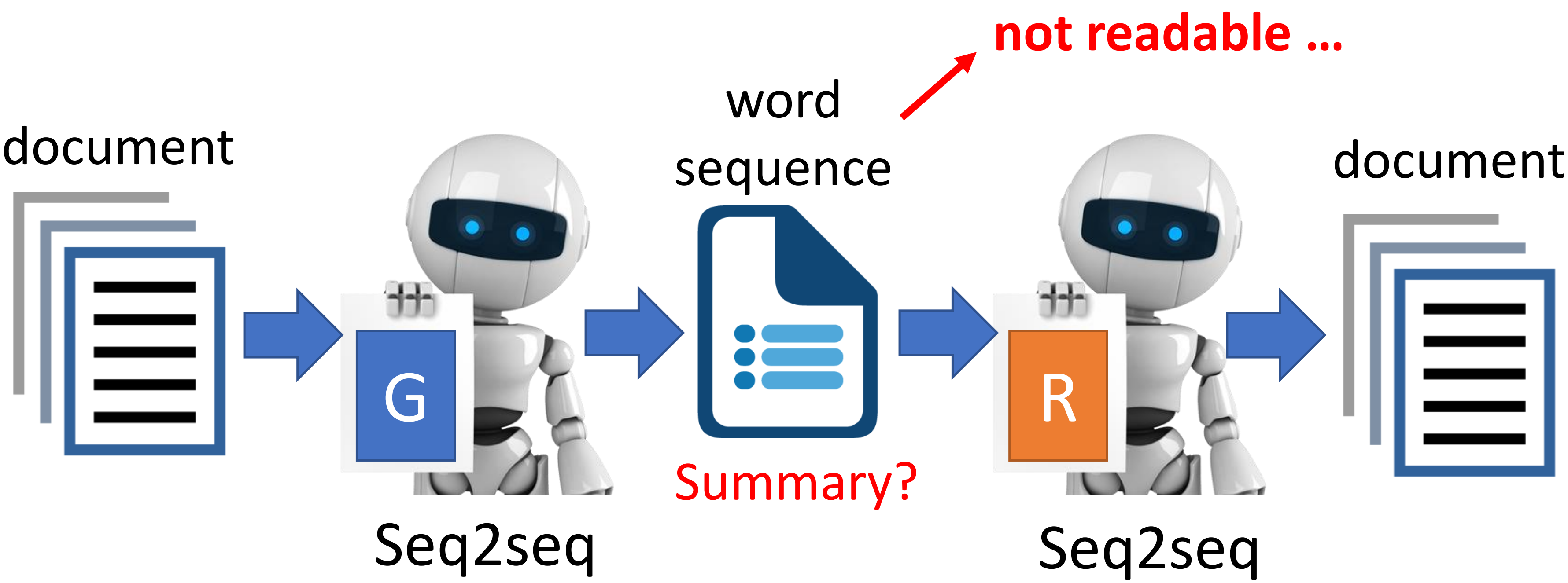


Unsupervised Abstractive Summarization

Only need a lot of documents to train the model

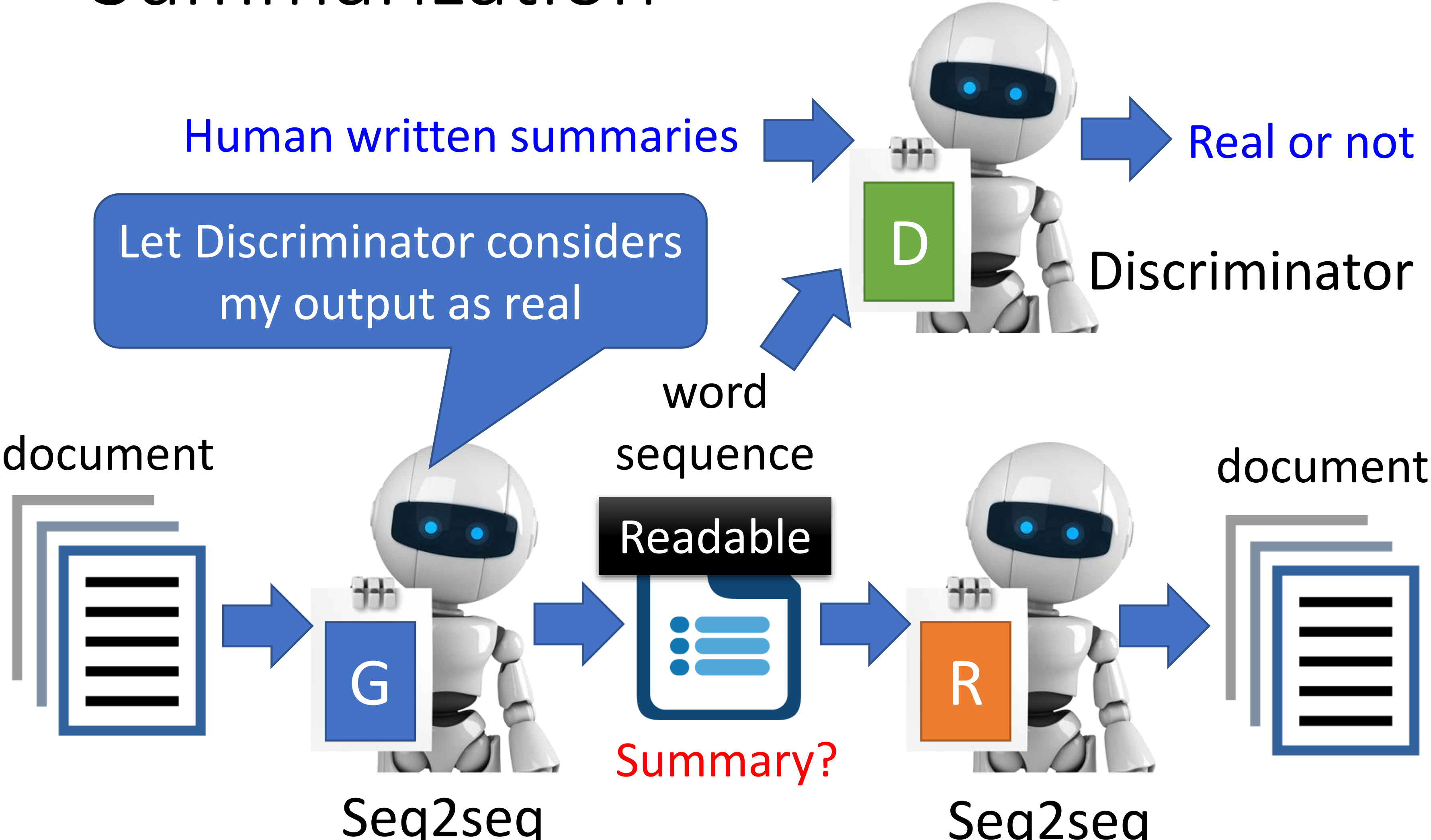
This is a ***seq2seq2seq auto-encoder***.

Using a sequence of words as latent representation.



Unsupervised Abstractive Summarization

REINFORCE algorithm is used.



Reference

- **Conditional Sequence Generation**

  - Jiwei Li, Will Monroe, Alan Ritter, Michel Galley, Jianfeng Gao, Dan Jurafsky, Deep Reinforcement Learning for Dialogue Generation, EMNLP, 2016
  - Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, Dan Jurafsky, Adversarial Learning for Neural Dialogue Generation, EMNLP, 2017
  - Matt J. Kusner, José Miguel Hernández-Lobato, GANS for Sequences of Discrete Elements with the Gumbel-softmax Distribution, arXiv 2016
  - Tong Che, Yanran Li, Ruixiang Zhang, R Devon Hjelm, Wenjie Li, Yangqiu Song, Yoshua Bengio, Maximum-Likelihood Augmented Discrete Generative Adversarial Networks, arXiv 2017
  - Lantao Yu, Weinan Zhang, Jun Wang, Yong Yu, SeqGAN: Sequence Generative Adversarial Nets with Policy Gradient, AAAI 2017

Reference

- **Conditional Sequence Generation**

  - Sai Rajeswar, Sandeep Subramanian, Francis Dutil, Christopher Pal, Aaron Courville, Adversarial Generation of Natural Language, arXiv, 2017
  - Ofir Press, Amir Bar, Ben Bogin, Jonathan Berant, Lior Wolf, Language Generation with Recurrent Generative Adversarial Networks without Pre-training, ICML workshop, 2017
  - Zhen Xu, Bingquan Liu, Baoxun Wang, Chengjie Sun, Xiaolong Wang, Zhuoran Wang, Chao Qi , Neural Response Generation via GAN with an Approximate Embedding Layer, EMNLP, 2017
  - Alex Lamb, Anirudh Goyal, Ying Zhang, Saizheng Zhang, Aaron Courville, Yoshua Bengio, Professor Forcing: A New Algorithm for Training Recurrent Networks, NIPS, 2016
  - Yizhe Zhang, Zhe Gan, Kai Fan, Zhi Chen, Ricardo Henao, Dinghan Shen, Lawrence Carin, Adversarial Feature Matching for Text Generation, ICML, 2017
  - Jiaxian Guo, Sidi Lu, Han Cai, Weinan Zhang, Yong Yu, Jun Wang, Long Text Generation via Adversarial Training with Leaked Information, AAAI, 2018
  - Kevin Lin, Dianqi Li, Xiaodong He, Zhengyou Zhang, Ming-Ting Sun, Adversarial Ranking for Language Generation, NIPS, 2017
  - William Fedus, Ian Goodfellow, Andrew M. Dai, MaskGAN: Better Text Generation via Filling in the_______, ICLR, 2018

Reference

- **Conditional Sequence Generation**

  - Sidi Lu, Yaoming Zhu, Weinan Zhang, Jun Wang, Yong Yu, Neural Text Generation: Past, Present and Beyond, arXiv, 2018
  - Yaoming Zhu, Sidi Lu, Lei Zheng, Jiaxian Guo, Weinan Zhang, Jun Wang, Yong Yu, Texygen: A Benchmarking Platform for Text Generation Models, arXiv, 2018
  - Zhen Yang, Wei Chen, Feng Wang, Bo Xu, Improving Neural Machine Translation with Conditional Sequence Generative Adversarial Nets, NAACL, 2018
  - Lijun Wu, Yingce Xia, Li Zhao, Fei Tian, Tao Qin, Jianhuang Lai, Tie-Yan Liu, Adversarial Neural Machine Translation, arXiv 2017
  - Linqing Liu, Yao Lu, Min Yang, Qiang Qu, Jia Zhu, Hongyan Li, Generative Adversarial Network for Abstractive Text Summarization, AAAI 2018
  - Rakshith Shetty, Marcus Rohrbach, Lisa Anne Hendricks, Mario Fritz, Bernt Schiele, Speaking the Same Language: Matching Machine to Human Captions by Adversarial Training, ICCV 2017
  - Xiaodan Liang, Zhiting Hu, Hao Zhang, Chuang Gan, Eric P. Xing, Recurrent Topic-Transition GAN for Visual Paragraph Generation, arXiv 2017

Reference

- **Text Style Transfer**

  - Zhenxin Fu, Xiaoye Tan, Nanyun Peng, Dongyan Zhao, Rui Yan, Style Transfer in Text: Exploration and Evaluation, AAAI, 2018
  - Tianxiao Shen, Tao Lei, Regina Barzilay, and Tommi Jaakkola, Style Transfer from Non-Parallel Text by Cross-Alignment, NIPS 2017
  - Chih-Wei Lee, Yau-Shian Wang, Tsung-Yuan Hsu, Kuan-Yu Chen, Hung-Yi Lee, Lin-shan Lee, Scalable Sentiment for Sequence-to-sequence Chatbot Response with Performance Analysis, ICASSP, 2018
  - Junbo (Jake) Zhao, Yoon Kim, Kelly Zhang, Alexander M. Rush, Yann LeCun, Adversarially Regularized Autoencoders, arxiv, 2017

Reference

- **Unsupervised Machine Translation**

  - Alexis Conneau, Guillaume Lample, Marc'Aurelio Ranzato, Ludovic Denoyer, Hervé Jégou, Word Translation Without Parallel Data, ICRL 2018
  - Guillaume Lample, Ludovic Denoyer, Marc'Aurelio Ranzato, Unsupervised Machine Translation Using Monolingual Corpora Only, ICRL 2018

- **Unsupervised Speech Recognition**

  - Da-Rong Liu, Kuan-Yu Chen, Hung-Yi Lee, Lin-shan Lee, Completely Unsupervised Phoneme Recognition by Adversarially Learning Mapping Relationships from Audio Embeddings, arXiv, 2018
  - Yi-Chen Chen, Chia-Hao Shen, Sung-Feng Huang, Hung-yi Lee, Towards Unsupervised Automatic Speech Recognition Trained by Unaligned Speech and Text only, arXiv, 2018
