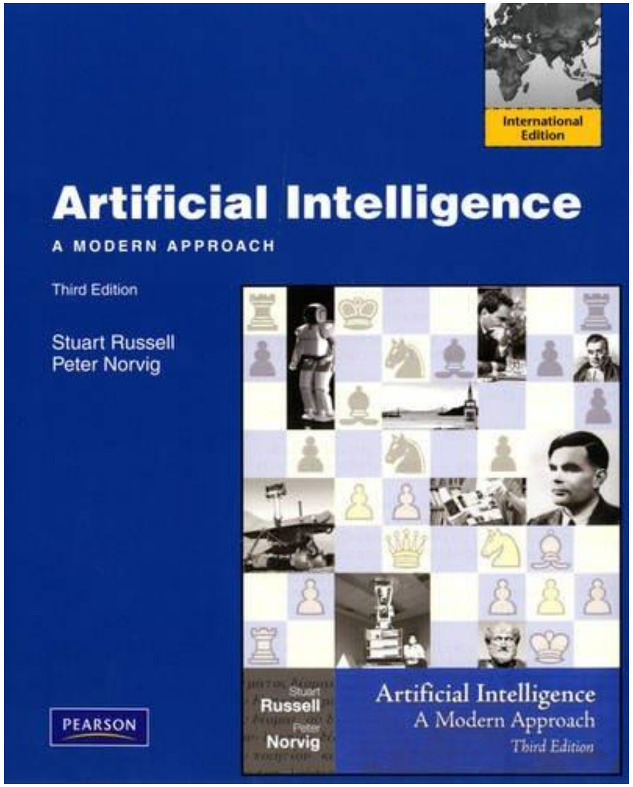
Intelligent Machines That Act Rationally

Hang Li
Toutiao AI Lab

Four Definitions of Artificial Intelligence

Building intelligent machines
(i.e., intelligent computers)

- Thinking humanly
- Acting humanly
- Thinking rationally
- Acting rationally



The Four Approaches to AI

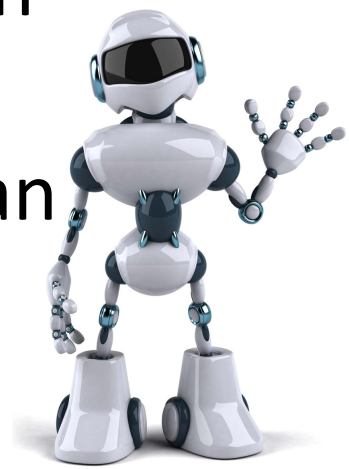
- Four approaches represent different motivations and views
- Four approaches need different theoretical foundations and methodologies
- If the goal is to build intelligent tools for humans, then taking the rational act approach is most appropriate
- Computers that act rationally also ‘think rationally’, but act is key
- We consider the rational act approach here

Talk Outline

- *Intelligent Machines That Act Rationally*
- Humans' Rational Behavior
- Architecture of Rationally Acting Machines
- Reinforcement Learning
- Our Work: Learning for Paraphrase Generation
- Summary

Characteristics of Machines Acting Rationally

1. Interaction with environment: acting as agent in an environment
2. Goal-oriented: acting to achieve a goal
3. Evaluation criterion: to evaluate performance of goal achievement
4. Functionalism: given an input, generate an output, not to mimic humans
5. Performance: comparable to or better than humans



Notes on Computers Acting Rationally

- Doing better than humans in various tasks does not mean realization of human intelligence
- Easier to make progress than the other approaches
- Because of not enough understanding of human brains' mechanism
- Becoming main stream of AI research now
- Due to the advancement of machine learning technologies

Talk Outline

- Intelligent Machines That Act Rationally
- *Humans' Rational Behavior*
- Architecture of Rationally Acting Machines
- Reinforcement Learning
- Our Work: Learning for Paraphrase Generation
- Summary

Question: how do humans think and act rationally?

Rene Descartes' View on Rationality

- Dualism: body and mind are separated
- Rationalism: rational thinking should only rely on reason, not on sensibility
- Reason = power of mind to think logically

Rene Descartes

Emotion is Essential in Human Thinking

- Three levels of mind
- Reason, emotion, and body function jointly participate in thinking
- We are not necessarily thinking machines; we are feeling machines that think – Antonio Damasio

Reason, Decision Making

Emotion, Feeling

Body Function

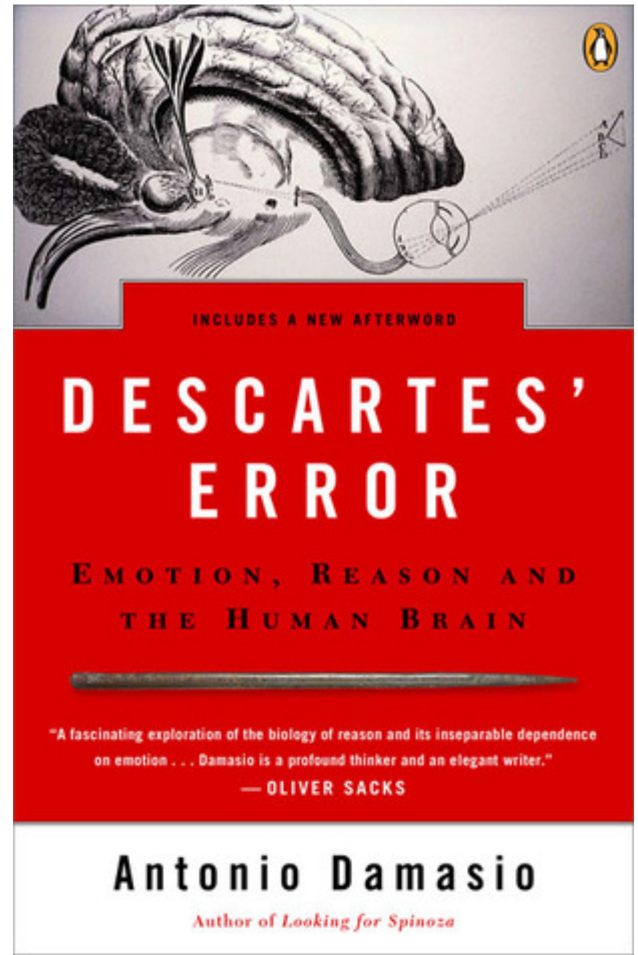
Experiment: Emotion and Judgement

- Setting
  - Subjects are divided into two groups
  - Taken into a room
  - First group is asked to recall experiences of social exclusion and second group is asked to recall experiences of social inclusion
  - Subjects are then asked to estimate temperature of room

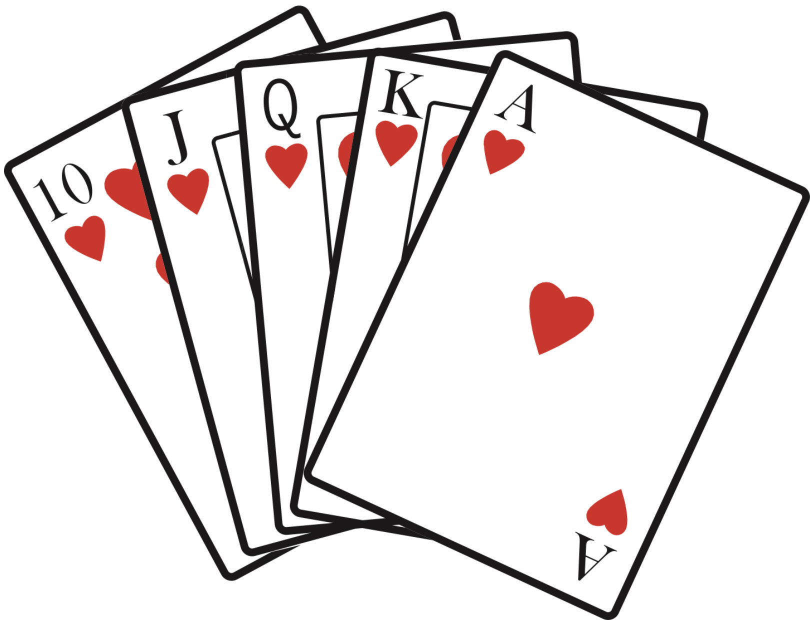
- Finding: average temperature of first group is lower than average temperature of second group
- Conclusion: emotion affects judgment

Rational Thinking Needs Emotion

- Descartes's dualism is wrong
- Rational think needs not only reason, but also emotion
- Patients who suffer from damages in emotion capability (obitofrontal cortex) cannot think and act rationally



Experiment: Emotion and Rationality



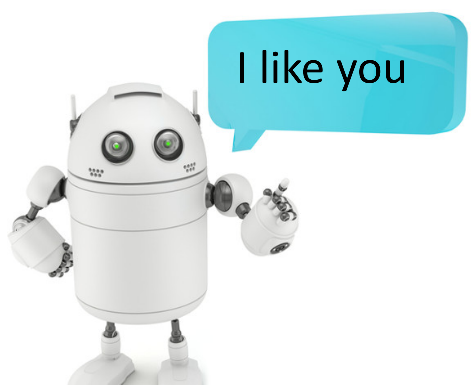
- Setting
  - Four decks of cards
  - Subject is asked to continuously pick up cards; depending on result, he can either win or lose money
  - Goal for subject is to win the game
  - Two safe decks: probability of winning is larger, amount of money in each win is small
  - Two risky decks: probability of losing is larger, amount of money in each win is large

Experiment: Emotion and Rationality

- Observation
  - Ordinary people pick up cards from risky decks at beginning, but gradually realize risk and move to safe decks
  - Patients who lost emotion capability do not feel pain of loss and continuously pick up cards from risky decks, until going bankrupt
- Conclusion
  - Emotion does affect rational behavior

To Build Computers Acting Rationally

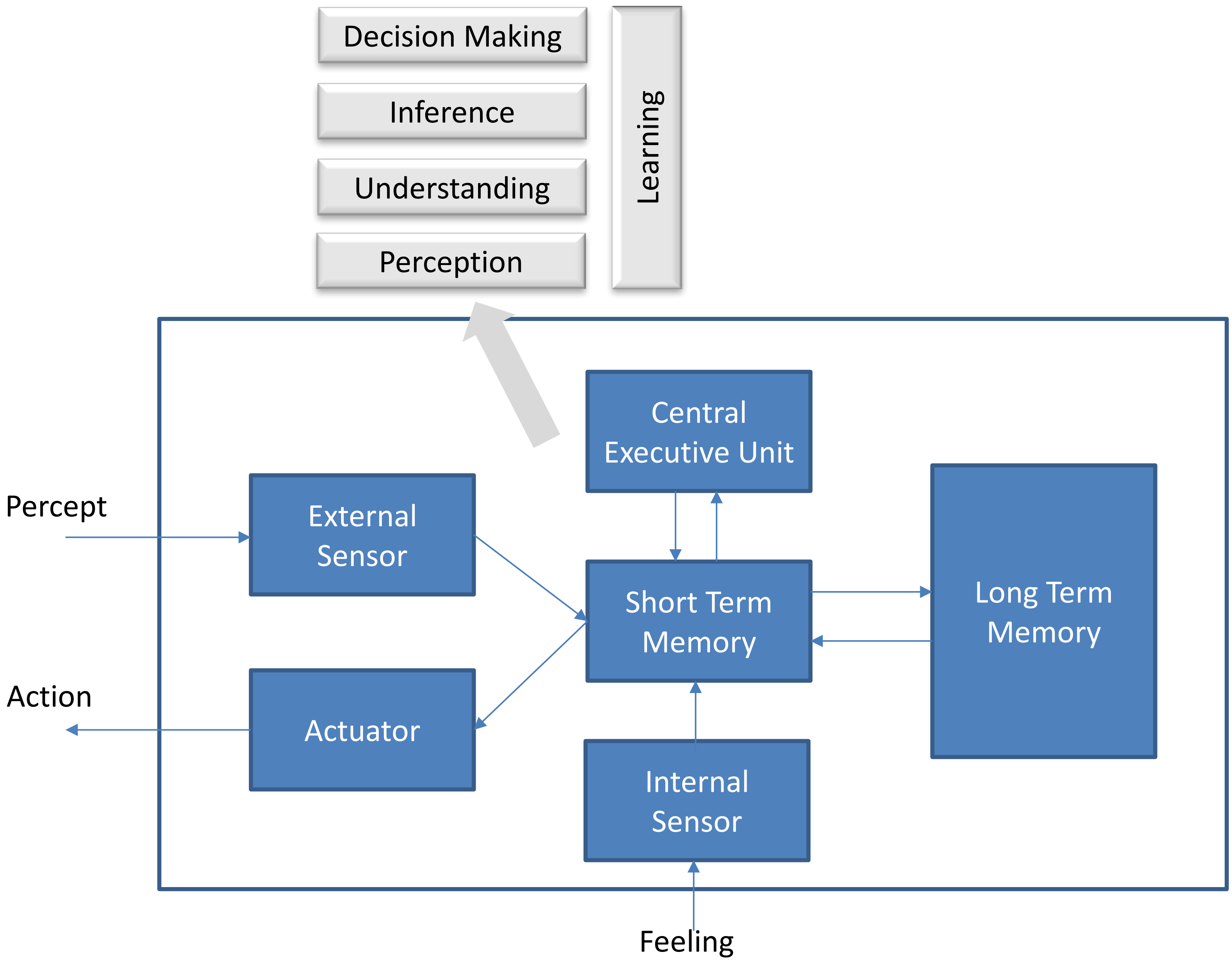
- Mainly include reason capability, and also emotion function and body function, depending on tasks
- Example: emotion is important factor for dialogue systems
- Example: self-charging is important feature for sweeping robots



Talk Outline

- Intelligent Machines That Act Rationally
- Humans' Rational Behavior
- *Architecture of Rationally Acting Machines*
- Reinforcement Learning
- Our Work: Learning for Paraphrase Generation
- Summary

An Architecture of Rationally Acting Machines



Notes on Architecture

- The architecture is defined at function level which should be shared by computers and humans
- Signals are from both environment and body
- Sensors transform inputs in different modalities into the same representations
- The use of long term memory and short term memory makes it possible to continuously acquire information and knowledge

Talk Outline

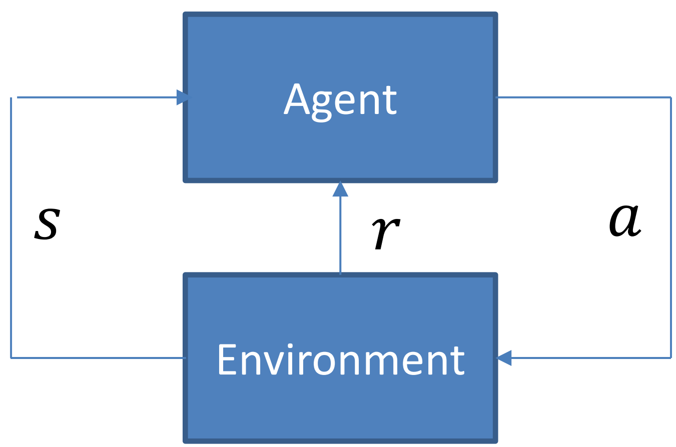
- Intelligent Machines That Act Rationally
- Humans' Rational Behavior
- Architecture of Rationally Acting Machines
- *Reinforcement Learning*
- Our Work: Learning for Paraphrase Generation
- Summary

Machine Learning for Rational Behavior

- Machine learning
  - Supervised learning
  - Unsupervised learning
  - Reinforcement learning
- Machine learning, particularly reinforcement learning, learns models for rational acts (recognition, understanding, inference, decision making)
- All formalized as optimization problems

Reinforcement Learning

- Data: $D = \{(s, a, r, s')\}_{t=0}^{T}$
- Model: $P(s'|s, a), R(s, a)$
- Policy function: $\pi(s)$: $\pi(s) = P(a|s)$ or $a = \pi(s)$
- Value function: $V(s)$: $E(\sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \,|\pi)$
- Model-based Learning: $P(s'|s, a), R(s, a) \rightarrow V(s) \rightarrow \pi(s)$
- Model-free Learning: $V(s) \rightarrow \pi(s)$ or $\pi(s)$



Reinforcement Learning

- Learned models are used in sequential decision process
- Reinforcement learning also includes partially observed Markov decision process, hierarchical reinforcement learning, inverse reinforcement learning
- Learning is formalized as optimization

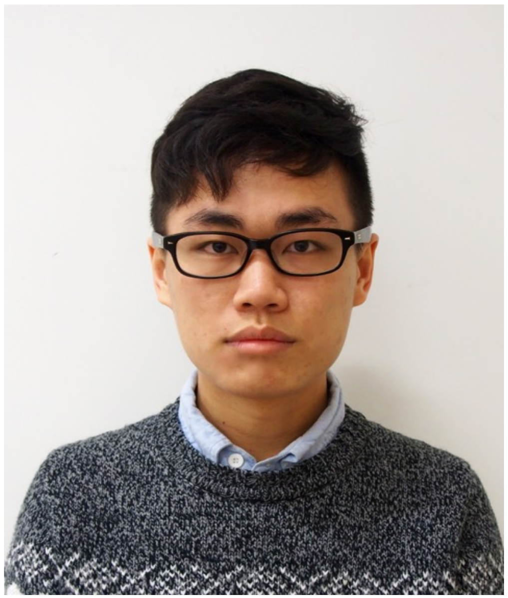
Talk Outline

- Intelligent Machines That Act Rationally
- Humans’ Rational Behavior
- Architecture of Rationally Acting Machines
- Reinforcement Learning
- *Our Work: Learning for Paraphrase Generation*
- Summary

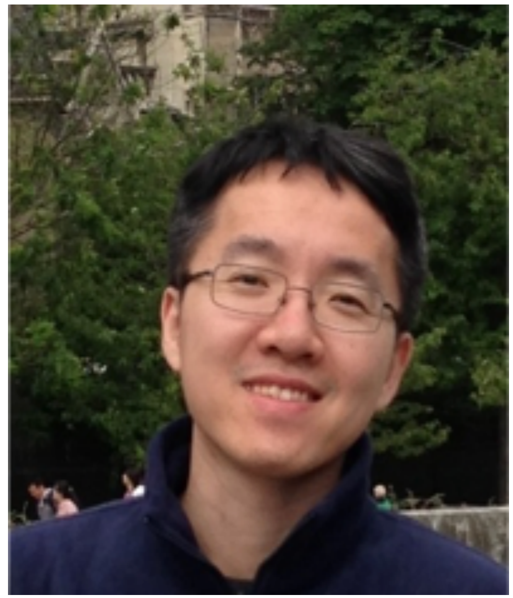
Not only learn how to act rationally but also learn how to evaluate rationality

(i.e., to learn both policy function and reward function in reinforcement learning)

Joint Work with



Zichao Li



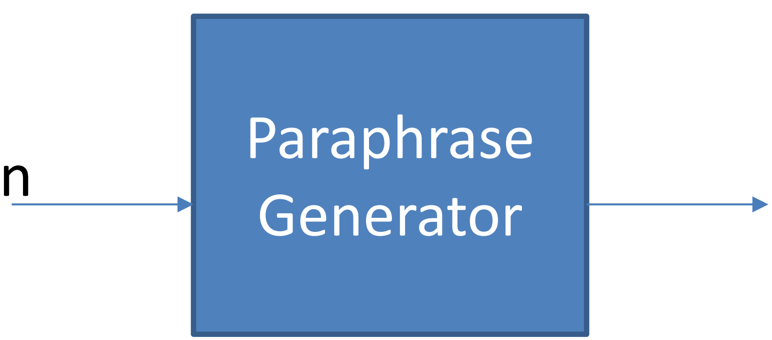
Xin Jiang

Lifeng Shang

Task: Paraphrase Generation

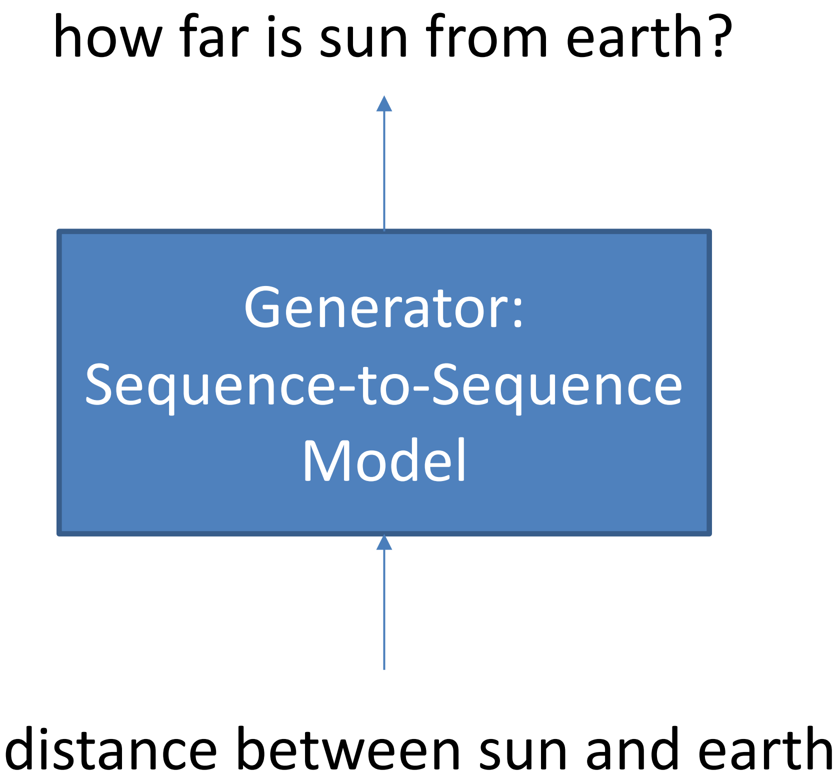
- Ambiguity and variability are nature of natural languages
- Ambiguity: same expression represents different meanings.
- Variability: different expressions represent same meaning
- Paraphrase auto generation: creates a number of synonymous expressions given an expression

- distance between sun and earth



- how far is sun from earth?
- how many miles is sun from earth?
- … …

Previous Work: Sequence to Sequence Model as Generator



- Supervised Learning
  - Training data: paraphrase pairs
  - Loss function: cross entropy
- Reinforcement Learning
  - Training data: paraphrase pairs
  - Reward: BLEU or ROUGE score

Main Challenge

- It is challenging to define a *semantic similarity measure* to guide the training of generator
- The evaluation measure needs to judgment whether two sentences are semantically similar
- Previously, lexical measures such as BLEU and ROUGE are used, which are not ideal

Ref: the Iraqi weapons are to be handed over to the army with two weeks
MT: in two weeks Iraq's weapons will give to the army

- 1-gram precision: 6/10
- 2-gram precision: 3/9
- 3-gram precision: 3/8

$$\mathrm{BLEU} = \left(\prod_{n=1}^{3} p_n\right)^{\frac{1}{3}} \quad \mathrm{BLEU} = \left(\frac{6}{10} * \frac{3}{9} * \frac{2}{8}\right)^{\frac{1}{3}} = 0.368$$

New Framework: Generator and Evaluator

- Using machine learning to learn both generator and evaluator
- Learning of evaluator is equally important

Evaluator:
Deep Matching Model

distance between sun and earth

Generator:
Sequence-to-Sequence Model

distance between earth and sun

Learning of Evaluator: Supervised Learning

1. Train evaluator with positive and negative examples
2. Train generator with positive examples and evaluator

(x_1^+, y_1^+)
(x_2^+, y_2^+)
……
(x_m^+, y_m^+)

Positive examples

(x_1^-, y_1^-)
(x_2^-, y_2^-)
……
(x_n^-, y_n^-)

Negative examples

Evaluator:
Deep Matching Model

(x_1^+, y_1^+)
(x_2^+, y_2^+)
……
(x_m^+, y_m^+)

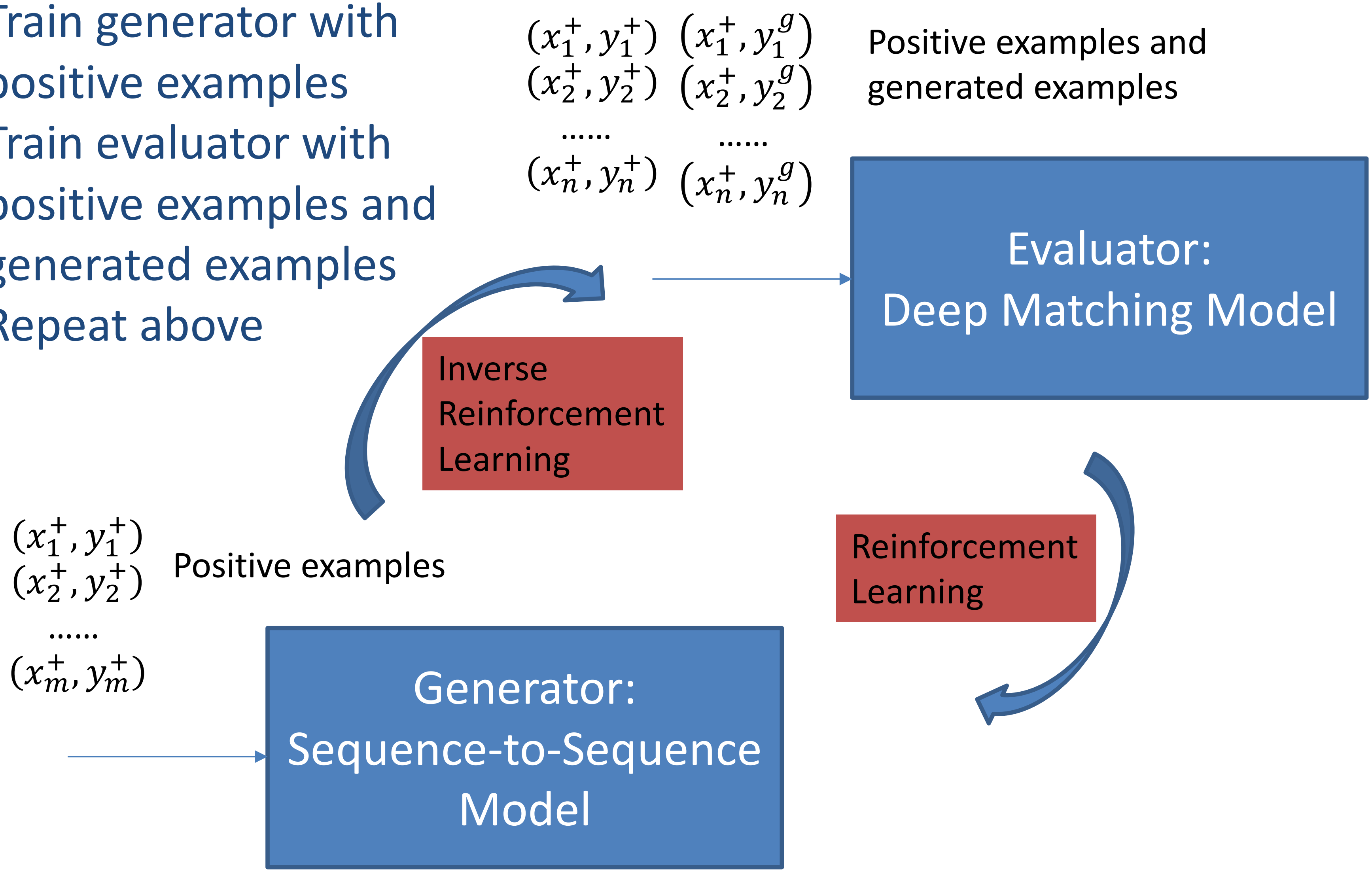
Positive examples

Generator:
Sequence-to-Sequence Model

Reinforcement Learning

Learning of Evaluator: Inverse Reinforcement Learning

1. Train generator with positive examples
2. Train evaluator with positive examples and generated examples
3. Repeat above



Inverse Reinforcement Learning

Find R^*, $\quad E(\sum_{t=0}^{\infty} \gamma^t R^*(s_t) | \pi^*) \geq E(\sum_{t=0}^{\infty} \gamma^t R^*(s_t) | \pi), \ \forall \pi$

Find M^*, $\quad E_{P(Y|X,\theta^*)} M^*(X,Y) \geq E_{P(Y|X,\theta)} M^*(X,Y), \forall \theta$

$$L_{IRL}(\phi) = E_{P(Y|X,\theta^*)} max\left(0, 1 + M_\phi(X,\hat{Y}) - M_\phi(X,Y) - \xi_X(\hat{Y},Y)\right)$$

Reinforcement Learning (Policy Gradient)

Find π^*, $\quad E(\sum_{t=0}^{\infty} \gamma^t R(s_t) | \pi^*) \geq E(\sum_{t=0}^{\infty} \gamma^t R(s_t) | \pi), \ \forall \pi$

Find θ^*, $\quad E_{P(Y|X,\theta^*)} M(X,Y) \geq E_{P(Y|X,\theta)} M(X,Y), \forall \theta$

$$\nabla_\theta L_{RL}(\theta) = E_{P(Y|X,\theta)} \nabla_\theta \log P(Y|X,\theta) M_\phi(X,Y)$$

Experimental Results

Model	ROUGE-1	BLEU-1	BLEU-2	METEOR
Generator: Seq2Seq	58.50	30.70	35.91	25.16
Generator: Seq2Seq with Copy	61.40	35.12	39.90	29.63
+Evaluator: Rouge	62.97	36.60	41.39	29.64
+Evaluator: SL	**63.69**	**37.61**	**42.85**	**31.78**
+Evaluator: IRL	63.63	37.33	42.68	31.62

Related Work

- AlphaGo
  - Monte Carlo Tree Search
  - Policy Network: to choose actions
  - Value Network: to evaluate actions
- GAN (Generative Adversarial Networks)
  - Generator: capturing data distribution
  - Discriminator: deciding whether data is from generator or true distribution

Talk Outline

- Intelligent Machines That Act Rationally
- Humans’ Rational Behavior
- Architecture of Rationally Acting Machines
- Reinforcement Learning
- Our Work: Learning for Paraphrase Generation
- *Summary*

Summary

- Machines that act rationally
  - Agent in environment, goal-oriented, evaluation criterion, functionalism, better performances
- Humans' rational behavior not only needs reason but also emotion
- Architecture of machines acting rationally
- Reinforcement learning is useful for machines acting rationally
- Framework for paraphrase generation, including learning of generator and evaluator

References

- Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, Third Edition, Prentice Hall, 2009.
- Anthony Damasio, Descartes' Error: Emotion, Reason, and the Human Brain, Penguin Books, 2005.
- Benjamin K. Bergen, Louder Than Words: The New Science of How the Mind Makes Meaning, Basic Books, 2012.
- Shum, He, and Li, From Eliza to Xiaoice, Challenges and Opportunities with Social Chatbots, 2018.
- Zichao Li, Xin Jiang, Lifeng Shang, Hang Li, Paraphrase Generation Using Deep Reinforcement Learning, arXiv:1711.00279, 2017.

Thank you!