Toward Neural Symbolic Processing

Hang Li
Noah’s Ark Lab
Huawei Technologies
Outline

• Overview on Natural Language Processing
• Neural Symbolic Processing
• Intelligent Information and Knowledge Management System
• Related Work
• Our Work
• Summary
Ultimate Goal: Natural Language Understanding

Natural Language Dialogue

Text Comprehension
Natural Language Understanding

• Two definitions:
  – Representation-based: if system creates proper internal representation, then we say it “understands” language
  – Behavior-based: if system properly follows instruction in natural language, then we say it “understands” language, e.g., “bring me a cup of tea”

• We take the latter definition
Five Characteristics of Human Language

• Incompletely Regular (Both Regular and Idiosyncratic)
• Compositional (or Recursive)
• Metaphorical
• Associated with Knowledge
• Interactive
Natural Language Understanding by Computer Is Extremely Difficult

• It is still not clear whether it is possible to realize human language ability on computer

• On modern computer
  – The incomplete regularity and compositionality characteristics imply complex combinatorial computation
  – The metaphor, knowledge, and interaction characteristics imply exhaustive computation

• Big question: can we invent new computer closer to human brain?
Reason of Challenge

• A computer system must be constructed based on math
• Open question: whether it is possible to process natural language as humans, using math models
• Natural language processing is believed to be AI complete
Question answering, including search, can be practically performed, because it is simplified.
Data-driven Approach May Work

• Hybrid is most realistic and effective for natural language processing, and AI
  – machine learning based
  – human-knowledge incorporated
  – human brain inspired

• Big data and deep learning provides new opportunity
Advancement in AI, including NLP can be made through the closed loop.
Fundamental Problems of Statistical Natural Language Processing

• Classification: assigning a label to a string
  \[ S \rightarrow C \]

• Matching: matching two strings
  \[ s, t \rightarrow R^+ \]

• Translation: transforming one string to another
  \[ S \rightarrow t \]

• Structured prediction: mapping string to structure given knowledge
  \[ S \rightarrow s' \]

• Sequential decision process: continuously choosing an action in a state, where the process randomly moves between states
Fundamental Problems of Statistical Natural Language Processing

- Classification
  - Text classification
  - Sentiment analysis

- Matching
  - Search
  - Question answering
  - Dialogue (single turn)

- Translation
  - Machine translation
  - Speech recognition
  - Handwriting recognition
  - Dialogue (single turn)

- Structured Prediction
  - Named entity extraction
  - Part of speech tagging
  - Sentence parsing
  - Semantic parsing

- Sequential Decision Process
  - Dialogue (multi turn, task dependent)
Lower Bound of User Need vs Upper Bound of Technology

Pushing Upper Bound of Technology
Future Trends of Natural Language Processing

• Speech recognition and machine translation are taking-off
  – There are still issues to be solved, e.g., long tail challenge
• Single turn dialogue and single turn question answering will take-off
  – Task-dependent single turn dialogue will be gradually used
  – Single turn question answering will be gradually used
• Multi-turn dialogue needs more research
  – Reinforcement learning can be key technology
  – Data needs to be collected first, and then the AI loop can be run
Outline

• Overview on Natural Language Processing
• Neural Symbolic Processing
• Intelligent Information and Knowledge Management System
• Related Work
• Our Work
• Summary
Combination of Neural Processing and Symbolic Processing

Symbolic Representation

- Easy to Interpret
- Easy to Manipulate

Neural Representation

- Able to Deal with Ambiguity
- Robust to Noise

Neural Symbolic Processing
Prof. Yoshua Bengio’s Comment

- Injecting symbols into neural works would be difficult, even impossible
- However, externally combining symbolic processing with neural processing should work
- This particularly makes sense for question answering and dialogue
Outline

• Overview on Natural Language Processing
• Neural Symbolic Processing
• Intelligent Information and Knowledge Management System
• Related Work
• Our Work
• Summary
Intelligent Information and Knowledge Management System

Use Phase

Q
A

Analysis

Language Processing Unit

Short-term Memory

Long-term Memory

Generation

Central Executive Unit

Retrieval
Intelligent Information and Knowledge Management System

Learning Phase

- Language Processing Unit
- Short-term Memory
- Long-term Memory
- Central Executive Unit

Q1 A1
Q2 A2
......
Characteristics and Current Status

• Continuously accumulates information and knowledge
• Properly performs question answering in natural language,
  – Answers when it knows
  – Says “I don’t know”, when it does not know
• Ideally, system is automatically constructed without human involvement
• Becomes intelligent assistant of human
  – Note that computer has two powerful capabilities, computing and storage
• Currently, only partially realized, cf., search engine
Neural Symbolic Processing for Information and Knowledge Management

Language Processing Model
- Encoder
- Decoder

Short-Term Memory
- Q
- A

Central Executive Unit

Long-Term Memory
- Sym.
- Neu.

Knowledge in symbolic representation & neural representation
Outline

• Overview on Natural Language Processing
• Neural Symbolic Processing
• Intelligent Information and Knowledge Management System
• Related Work
• Our Work
• Summary
Semantic Parsing

- **Executor**: execute command based on logic form and context \( y = [z]_c \)
- **Grammar**: set of rules for creating derivations based on input \( D(x, c) \) and context
- **Model**: model for ranking derivations with parameters \( P_\theta(d \mid x, c) \)
- **Parser**: find most likely derivation under learned model \( d^* \)
- **Learner**: learn parameters of model \( \theta \) from data \( \{(x_i, c_i, y_i)\}_{i=1}^{n} \)

**Q**: What is the largest prime less than 10?  
**A**: 7

Liang 2016
Never Ending Language Learning (NELL)

**Task**
- Initial ontology, few examples of each category predicate, the web, occasional interaction from humans
- Extract more facts from web
- Learn to read better than before

**System**
- KB with 15 million candidate beliefs

**Technologies**
- Coupled semi-supervised learning, automatic discovery of new coupling constraints, automatic extending of ontology, staged curriculum

Mitchell et al. 2015
Memory Networks

• Long term memory + inference
• Model is learned
• Can answer factoid questions
• Acc = 40%+

Example
− John is in the playground.
− Bob is in the office.
− John picked up the football.
− Bob went to the kitchen.
− Q: where is the football?
− A: playground

Weston et al. 2014
Differentiable Neural Computers

- DNC = neural network + external memory (matrix)
- Memory represents complex data structures
- Neural network, learned from data and supervised learning, controls access to memory
- Memory heads use three forms of differentiable attention
- Resembling mammalian hippocampus

Graves et al. 2016
Outline

• Overview on Natural Language Processing
• Neural Symbolic Processing
• Intelligent Information and Knowledge Management System
• Related Work
• Our Work
• Summary
Researchers

Zhengdong Lu  Xin Jiang  Lifeng Shang
Q: How tall is Yao Ming?
A: He is 2.29m tall and is visible from space.  
(Yao Ming, height, 2.29m)

Q: Which country was Beethoven from?  
A: He was born in what is now Germany.  
(Ludwig van Beethoven, place of birth, Germany)

Q: How tall is Liu Xiang?  
A: He is 1.89m tall
GenQA

Language Processing Unit

Encoder creates question representation, decoder generates answer

Short-Term Memory

Matches and retrieves most relevant answer representation

Long-Term Memory

Triples in symbolic representations (indexed) & neural representation

End-to-End Training
Q: How many people participated in the game in Beijing?
A: 4,200
SQL: `select #_participants, where city=beijing`

Q: When was the latest game hosted?
A: 2012
SQL: `argmax(city, year)`

Q: Which city hosted the longest Olympic game before the game in Beijing?
A: Athens
Neural Enquirer

Language Processing Unit

Encoder

Q

Decoder

A

Short-Term Memory

Matching

Q'

A'

Matching

Long-Term Memory

Feature and values are in symbolic representations and neural representation

Encoder creates question representation, decoder simply returns answer

Matches question representation to table representations to find answer
Summary

• Intelligent Information and Management System
  – Can automatically acquire information and knowledge
  – Can accurately answer questions from humans

• Should be most important topic for research in AI

• Neural Symbolic Processing should be most promising approach

• Recent research is making progress

• Many open questions and challenges
References

- 李航，迎接自然语言处理新时代，计算机学会通讯，2017年第2期
- 李航，简论人工智能，计算机学会通讯，2016年第3期
- 李航，对于AI我们应该期待什么，计算机学会通讯，2016月第11期
- 李航，技术的上界与需求的下界，新浪博客，2014年
Thank You!