From Symbolic to Neural Approaches to NLP  
- Case Studies of Machine Reading and Dialogue

Jianfeng Gao

Thanks for the slides by Bill Dolan, Michel Galley, Lihong Li, Yi-Min Wang et al.

Joint work with many Microsoft colleagues and interns (see the list of collaborators)

Microsoft AI & Research

Nov. 11, 2017, Dalian, China
Outline

• Part 1: The transition of NLP to neural approaches
  • Deep learning leads to paradigm shift in NLP
  • The powers of deep learning
  • Cast study of Deep Semantic Similarity Models

• Part 2: Neural machine reading models for question answering

• Part 3: Deep reinforcement learning for task-completion dialogue
Traditional definition of NLP: the branch of AI

• Deal with analyzing, understanding and generating the languages that humans use naturally (natural language)

• Study knowledge of language at different levels
  • Phonetics and Phonology – the study of linguistic sounds
  • Morphology – the study of the meaning of components of words
  • Syntax – the study of the structural relationships between words
  • Semantics – the study of meaning
  • Discourse – they study of linguistic units larger than a single utterance

[Jurafsky & Martin 10]
Traditional NLP component stack

1. **Natural language understand (NLU):** parsing (speech) input to semantic meaning and update the system state
2. **Application reasoning and execution:** take the next action based on state
3. **Natural language generation (NLG):** generating (speech) response from action

[Menezes & Dolan 17]
**Pragmatic definition:** building computer systems

- Process large text corpora, turning information into knowledge
  - Text classification
  - Information retrieval and extraction
  - Machine reading comprehension and question answering
  - ...
- Enable human-computer interactions, making knowledge accessible to humans in the most natural way
  - Dialogue and conversational agents
  - Machine translation
  - ...

Challenge of NLP: the diversity of natural language

Many-to-many mapping btw symbolic language and semantic meaning

Ambiguity
Example: I made her duck.
- I cooked waterfowl for her.
- I cooked waterfowl belonging to her.
- I created the plaster duck she owns.
- I caused her to quickly lower her head or body.
- I waved my magic wand and turned her into undifferentiated waterfowl.

Paraphrase
Example: How long is the X river?
- The Mississippi River is 3,734 km (2,320 mi) long.
- ...is a short river, some 4.5 miles (7.2 km) in length
- The total length of the river is 2,145 kilometers.
- ... at the estimated length of 5,464 km (3,395 mi)...
- ... has a meander length of 444 miles (715 km)...
- ... Bali’s longest river, measuring approximately 75 kilometers from source to mouth.
- The ... mainstem is 2.75 miles (4.43 km) long although total distance from headwater source tributaries to the sea is 14 miles (23 km).

[Jurafsky & Martin 10; Dolan 17]
Deep Learning (DL) leads to a paradigm shift in NLP: from symbolic to neural approaches

<table>
<thead>
<tr>
<th>Traditional symbolic approaches</th>
<th>Neural approaches</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Discrete, symbolic space</td>
<td>• Continuous, neural space</td>
</tr>
<tr>
<td>• Human comprehensible</td>
<td>• Human incomprehensible</td>
</tr>
<tr>
<td>• easy to debug</td>
<td>• hard to debug</td>
</tr>
<tr>
<td>• Computationally inefficient</td>
<td>• Computationally efficient</td>
</tr>
<tr>
<td>• Sensitive to ambiguity/paraphrase</td>
<td>• Robust to ambiguity/paraphrase</td>
</tr>
<tr>
<td>• Cascaded models prone to error propagation and require careful feature engineering</td>
<td>• E2E learning leads to better performance and simplified systems</td>
</tr>
</tbody>
</table>
E2E approaches based on DL

Discrete, symbolic space
- Human comprehensible
- Input: $x$
- Output: $y$

Continuous, neural space
- Computationally efficient
- Input: $h_x$
- Output: $h_y$

$x = f_e(h_x; \theta_e)$, Symbolic $\rightarrow$ Neural by embedding models / encoder

$y = f_d(h_y; \theta_d)$, Neural $\rightarrow$ Symbolic by generative models / decoder

$L(\theta) \propto \text{Error}(y, y^*)$
The powers of deep learning

1. **End-to-end Learning**
   - Simplifies systems, reduces effort for feature engineering and localization

2. **Strong Representation Power**
   - Due to novel DNN architectures and learning algorithms; leads to high accuracy in many tasks

3. **Semantic Representation Learning**
   - Leads to a paradigm shift in NLP/IR: from symbolic to neural computation

4. **New Applications and Experience**
   - E.g., link language to real-world signals such as images and machine state

5. **Deep Reinforcement Learning**
   - Makes it possible to build intelligent agents for real-world applications such as goal-oriented dialogue

### State of the art results on NLP application-level tasks

<table>
<thead>
<tr>
<th>Task</th>
<th>Test set</th>
<th>Metric</th>
<th>Best non-neural</th>
<th>Best neural</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Machine Translation</td>
<td>Enu-deu newstest16</td>
<td>BLEU</td>
<td>31.4</td>
<td>34.8</td>
<td><a href="http://matrix.statmt.org">http://matrix.statmt.org</a></td>
</tr>
<tr>
<td></td>
<td>Deu-enu newstest16</td>
<td>BLEU</td>
<td>35.9</td>
<td>39.9</td>
<td><a href="http://matrix.statmt.org">http://matrix.statmt.org</a></td>
</tr>
<tr>
<td>Sentiment Analysis</td>
<td>Stanford sentiment bank</td>
<td>5-class Accuracy</td>
<td>71.0</td>
<td>80.7</td>
<td>Socher+ 13</td>
</tr>
<tr>
<td>Question Answering</td>
<td>WebQuestions test set</td>
<td>F1</td>
<td>39.9</td>
<td>52.5</td>
<td>Yih+ 15</td>
</tr>
<tr>
<td>Entity Linking</td>
<td>Bing Query Entity Linking set</td>
<td>AUC</td>
<td>72.3</td>
<td>78.2</td>
<td>Gao+ 14b</td>
</tr>
<tr>
<td>Image Captioning</td>
<td>COCO 2015 challenge</td>
<td>Turing test pass%</td>
<td>25.5</td>
<td>32.2</td>
<td>Fang+ 15</td>
</tr>
<tr>
<td>Sentence compression</td>
<td>Google 10K dataset</td>
<td>F1</td>
<td>0.75</td>
<td>0.82</td>
<td>Fillipova+ 15</td>
</tr>
<tr>
<td>Response Generation</td>
<td>Sordoni dataset</td>
<td>BLEU-4</td>
<td>3.98</td>
<td>5.82</td>
<td>Li+ 16a</td>
</tr>
</tbody>
</table>

[Menezes & Dolan 17]
Deep Semantic Similarity Model (DSSM)

- Compute semantic similarity between two text strings X and Y
  - Map X and Y to feature vectors in a latent semantic space via deep neural net
  - Compute the cosine similarity between the feature vectors

<table>
<thead>
<tr>
<th>Tasks</th>
<th>X</th>
<th>Y</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web search</td>
<td>Search query</td>
<td>Web document</td>
<td>Huang+ 13; Shen+ 14; Palangi+ 16</td>
</tr>
<tr>
<td>Entity linking</td>
<td>Entity mention and context</td>
<td>Entity and its corresponding page</td>
<td>Gao+ 14b</td>
</tr>
<tr>
<td>Online recommendation</td>
<td>Doc in reading</td>
<td>Interesting things / other docs</td>
<td>Gao+ 14b</td>
</tr>
<tr>
<td>Image captioning</td>
<td>Image</td>
<td>Text</td>
<td>Fang+ 15</td>
</tr>
<tr>
<td>Machine translation</td>
<td>Sentence in language A</td>
<td>Translations in language B</td>
<td>Gao+ 14a</td>
</tr>
<tr>
<td>Question answering</td>
<td>Question</td>
<td>Answer</td>
<td>Yih+ 15</td>
</tr>
</tbody>
</table>
3. Semantic Representation Learning

- Fast food
- Hot dog
- Dog racing
DSSM for entity linking

The Einstein Theory of Relativity

(1) The perihelion of Mercury shows a discrepancy which has long puzzled astronomers. This discrepancy is fully accounted for by Einstein. At the time when he published his theory, this was its only experimental verification.

(2) Modern physicists were willing to suppose that light might be subject to gravitation—i.e., that a ray of light passing near a great mass like the sun might be deflected to the extent to which a particle moving with the same velocity would be deflected according to the orthodox theory of gravitation. But Einstein's theory required that the light should be deflected just twice as much as this. The matter could only be tested during an eclipse among a number of bright stars. Fortunately a peculiarly favourable eclipse occurred last year. The results of the observations
DSSM: Compute Similarity in Semantic Space

Relevance measured by cosine similarity

Learning: maximize the similarity between X (source) and Y (target)

Representation: use DNN to extract abstract semantic features, $f$ or $g$ is a
- Multi-Layer Perceptron (MLP) if text is a bag of words [Huang+ 13]
- Convolutional Neural Network (CNN) if text is a bag of chunks [Shen+ 14]
- Recurrent Neural Network (RNN) if text is a sequence of words [Palangi+ 16]
DSSM: Compute Similarity in Semantic Space

Relevance measured by cosine similarity

Semantic layer $h$
Max pooling layer $v$
Convolutional layer $c_t$
Word hashing layer $f_i$
Word sequence $x_t$

Learning: maximize the similarity between X (source) and Y (target)

Representation: use DNN to extract abstract semantic representations

Convolutional and Max-pooling layer: identify key words/concepts in X and Y

Word hashing: use sub-word unit (e.g., letter $n$-gram) as raw input to handle very large vocabulary
Learning DSSM from Labeled X-Y Pairs

- Positive X-Y pairs are extracted from search click logs
- Negative X-Y pairs are randomly sampled
- Map X and Y into the same semantic space via deep neural net
Learning DSSM from Labeled X-Y Pairs

- Positive X-Y pairs are extracted from search click logs
- Negative X-Y pairs are randomly sampled
- Map X and Y into the same semantic space via deep neural net
- Positive Y are closer to X than negative Y in that space
Learning DSSM from Labeled X-Y Pairs

• Consider a query $X$ and two docs $Y^+$ and $Y^-$
  • Assume $Y^+$ is more relevant than $Y^-$ with respect to $X$
• $\text{sim}_\theta(X, Y)$ is the cosine similarity of $X$ and $Y$ in semantic space, mapped by DSSM parameterized by $\theta$

• $\Delta = \text{sim}_\theta(X, Y^+) - \text{sim}_\theta(X, Y^-)$
  • We want to maximize $\Delta$
• $\text{Loss}(\Delta; \theta) = \log(1 + \exp(-\gamma \Delta))$
• Optimize $\theta$ using mini-batch SGD on GPU
New Applications and Experience

Neural approaches allow language models to be grounded in the world, i.e., link language to real-world signals such as images, machine state, sensor data from biomedical devices.

Output of a neural conversation model trained on 250K Twitter conversations sparked by a tweeted photo

[Menezes & Dolan 17; Sordoni+ 15; Li+ 16a; MSR Data-Driven Conversation]
Social Bots [MSR Data-Driven Conversation]

• The success of XiaoIce (小冰)

• Problem setting and evaluation
  • Maximize the user engagement by automatically generating
  • *enjoyable* and *useful* conversations

• Learning a neural conversation engine
  • A data driven engine trained on social chitchat data [Sordoni+ 15; Li+ 16a]
  • Persona based models and speaker-role based models [Li+ 16b; Luan+ 17]
  • Image-grounded models [Mostafazadeh+ 17]
  • Knowledge-grounded models [Ghazvininejad+ 17]
Outline

• Part 1: The transition of NLP to neural approaches

• Part 2: Neural machine reading models for question answering
  • MindNet: a case study of symbolic approaches
  • Neural approaches to MRC and QA
  • ReasoNet: a case study of neural approaches
  • Ongoing research: visualize the reasoning process in neural space

• Part 3: Deep reinforcement learning for task-completion dialogue
Question Answering (QA) on Knowledge Base

Large-scale knowledge graphs
• Properties of billions of entities
• Plus relations among them

An QA Example:

**Question:** what is Obama’s citizenship?
• Query parsing:
  (Obama, Citizenship, ?)
• Identify and infer over relevant subgraphs:
  (Obama, BornIn, Hawaii)
  (Hawaii, PartOf, USA)
• correlating semantically relevant relations:
  BornIn ~ Citizenship
**Answer:** USA
Symbolic approaches to QA: production system

https://en.wikipedia.org/wiki/Production_system_(computer_science)

• Production rules
  • condition—action pairs
  • Represent (world) knowledge as a graph

• Working memory
  • Contains a description of the current state of the world in a reasoning process

• Recognizer-act controller
  • Update working memory by searching and firing a production rule

• A case study: MSR MindNet [Dolan+ 93; Richardson+ 98]
Case study of Question Answering with MindNet

• Build a MindNet graph from:
  • Text of dictionaries
  • Target corpus, e.g. an encyclopedia (Encarta 98)

• Build a dependency graph from query

• Model QA as a graph matching procedure
  • Heuristic fuzzy matching for synonyms, named entities, wh-words, etc.
  • Some common sense reasoning (e.g. dates, math)

• Generate answer string from matched subgraph
  • Including well-formed answers that didn’t occur in original corpus
Fragment of lexical space surrounding “bird”
Fuzzy Match against MindNet

Input LF:

Who assassinated Abraham Lincoln?

American actor John Wilkes Booth, who was a violent backer of the South during the Civil War, shot Abraham Lincoln at Ford's Theater in Washington, D.C., on April 14, 1865.
Generate output string

“John Wilkes Booth shot Abraham Lincoln”
Worked beautifully!

• Just not very often…
• What went wrong?
  • One major reason: paraphrase alternations

  • The Mississippi River is 3,734 km (2,320 mi) long.
  • ...is nearly 86 km long…
  • ...is a short river, some 4.5 miles (7.2 km) in length
  • The total length of the river is 2,145 kilometres (1,333 mi).
  • ... at the estimated length of 5,464 km (3,395 mi)...
  • ...is a 25-mile (40 km) tributary of ...
  • ... has a meander length of 444 miles (715 km)...
  • ... Bali’s longest river, measuring approximately 75 kilometers from source to mouth.
  • The ... mainstem is 2.75 miles (4.43 km) long although total distance from headwater source tributaries to the sea is 14 miles (23 km).
### Symbolic Space

- **Knowledge Representation**
  - *Explicitly* store a big but incomplete knowledge graph (KG)
  - Words, relations, templates
  - High-dim, discrete, sparse vectors

- **Inference**
  - Slow on a big KG
  - Keyword/template matching is sensitive to paraphrase alternations

- **Human comprehensible but not computationally efficient**

### Neural Space

- **Knowledge Representation**
  - *Implicitly* store entities and structure of KG in a *compact* way that is *more generalizable*
  - Semantic concepts/classes
  - Low-dim, cont., dense vectors shaped by KG

- **Inference**
  - Fast on compact memory
  - Semantic matching is *robust* to paraphrase alternations

- **Computationally efficient but not human comprehensible yet**

---

*Image of a knowledge graph and semantic concepts:*

```
“film”, “award”
film-genre/films-in-this-genre
film/cinematography
cinematographer/film
award-honor/honored-for
netflix-title/netflix-genres
director/film
award-honor/honored-for
```
From symbolic to neural computation

Symbolic Space
- UI: human readable I/O
- Leverage traditional symbolic approaches as pre/post processing
  - Keyword matching
  - Ontology based models
  - e.g., doc/passage/entity search/ranking

Neural Space

Input: Q

Question: Symbolic $\rightarrow$ Neural by embedding models / encoder

Error(A, A*)

Output: A

Answer: Neural $\rightarrow$ Symbolic by generative models / decoder

Inference: Question + KG $\rightarrow$ Answer by RNN + Attention
Case study: ReasoNet with Shared Memory

- **Production Rules** $\rightarrow$ **Shared memory** encodes task-specific knowledge
  - **Long-term memory**: encode KB for answering all questions in QA on KB
  - **Short-term memory**: encode the passage(s) which contains the answer of a question in QA on Text

- **Working memory** $\rightarrow$ **Hidden state** $S_t$ Contains a description of the current state of the world in a reasoning process

- **Recognizer-act controller** $\rightarrow$ **Search controller** performs multi-step inference to update $S_t$ of a question using knowledge in shared memory

- Input/output modules are task-specific

[Shen+ 16a]
KB relation paths in symbolic vs. neural spaces

Symbolic Space

Neural Space

[Shen+ 16a]
Search controller for KB QA

[Shen+ 16a]
Joint learning of Shared Memory and Search Controller

Training samples from KG:
(Obama, BornIn, Hawaii)
(Hawaii, PartOf, USA)
...
(h, r, t)
...
(Obama, Citizenship, ?) -> (USA)
Training samples from KG:
(Obama, BornIn, Hawaii)
(Hawaii, PartOf, USA)
...
(h, r, t)
...
(Obama, Citizenship, ?) -> (USA)
Shared Memory: long-term memory to store learned knowledge, like human brain

- Knowledge is learned via performing tasks, e.g., update memory to answer new questions
- New knowledge is *implicitly* stored in memory cells via gradient update
- Semantically relevant relations/entities can be compactly represented using similar vectors.
The Knowledge Base Question Answering Results on WN18 and FB15K

<table>
<thead>
<tr>
<th>Model</th>
<th>Additional Information</th>
<th>WN18</th>
<th>FB15k</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Hits@10 (%)</td>
<td>MR</td>
</tr>
<tr>
<td>SE (Bordes et al., 2011)</td>
<td>NO</td>
<td>80.5</td>
<td>985</td>
</tr>
<tr>
<td>Unstructured (Bordes et al., 2014)</td>
<td>NO</td>
<td>38.2</td>
<td>304</td>
</tr>
<tr>
<td>TransE (Bordes et al., 2013)</td>
<td>NO</td>
<td>89.2</td>
<td>251</td>
</tr>
<tr>
<td>TransH (Wang et al., 2014)</td>
<td>NO</td>
<td>86.7</td>
<td>303</td>
</tr>
<tr>
<td>TransR (Lin et al., 2015b)</td>
<td>NO</td>
<td>92.0</td>
<td>225</td>
</tr>
<tr>
<td>CTransR (Lin et al., 2015b)</td>
<td>NO</td>
<td>92.3</td>
<td>218</td>
</tr>
<tr>
<td>KG2E (He et al., 2015)</td>
<td>NO</td>
<td>93.2</td>
<td>348</td>
</tr>
<tr>
<td>TransD (Ji et al., 2015)</td>
<td>NO</td>
<td>92.2</td>
<td>212</td>
</tr>
<tr>
<td>TATEC (García-Durán et al., 2015)</td>
<td>NO</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NTN (Socher et al., 2013)</td>
<td>NO</td>
<td>66.1</td>
<td>-</td>
</tr>
<tr>
<td>DISTMULT (Yang et al., 2014)</td>
<td>NO</td>
<td>94.2</td>
<td>-</td>
</tr>
<tr>
<td>STTransE (Nguyen et al., 2016)</td>
<td>NO</td>
<td>94.7 (93)</td>
<td>244 (206)</td>
</tr>
<tr>
<td>RTransE (García-Durán et al., 2015)</td>
<td>Path</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PTransE (Lin et al., 2015a)</td>
<td>Path</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>NLFeat (Toutanova et al., 2015)</td>
<td>Node + Link Features</td>
<td>94.3</td>
<td>-</td>
</tr>
<tr>
<td>Random Walk (Wei et al., 2016)</td>
<td>Path</td>
<td>94.8</td>
<td>-</td>
</tr>
<tr>
<td>ReasoNet (Shen+ 16a)</td>
<td>NO</td>
<td>95.3</td>
<td>249</td>
</tr>
</tbody>
</table>
Visualization of Reasoning in MRC Models

- Translate Natural Language to Image
- Multi-step Image Editing via Dialogue

“A living room”

“A white sofa by the green wall”

“There is a blue vase by the foot of the sofa”
First Step: Let there be Color, and Shape!

Task 1: **Text-guided Image Colorization**

![BW Image](image1.png) + "The flower has purple petals with a white stamen" → ![Color Image](image2.png)

Task 2: **Text-guided Image Segmentation**

![RGB-depth Image](image3.png) + "At the middle of the kitchen lies a blue chair. The upper cabinet has a black microwave. There is a black garbage can on the left of the blue chair. There is a white garbage can on the right of the blue chair. On its left there is a red fire distinguisher." → ![Segmentation](image4.png)
Multi-Modal ReasoNet

The flower has purple petals with a white stamen

Multimodal ReasoNet
Outline

• Part 1: The transition of NLP to neural approaches
• Part 2: Neural machine reading models for question answering

• **Part 3: Deep reinforcement learning for task-completion dialogue**
  • Dialogue as RL
  • Case study 1: InfoBot with end-to-end learning RL
  • Case study 2: Composite task completion bot with Hierarchical RL
  • Ongoing research: subgoal discovery for hierarchical RL
Multi-turn (goal-oriented) dialogue

“Find me a Bill Murray movie”

(Spoken) Language Understanding

“When was it released”

Natural Language Generation / Synthesis

Dialog Manager

State Tracking

Request(movie; actor=bill murray)

Request (release_year)

Dialog Policy

Knowledge Base

<table>
<thead>
<tr>
<th>Movie</th>
<th>Actor</th>
<th>Release Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groundhog Day</td>
<td>Bill Murray</td>
<td>1993</td>
</tr>
<tr>
<td>Australia</td>
<td>Nicole Kidman</td>
<td>X</td>
</tr>
<tr>
<td>Mad Max: Fury Road</td>
<td>X</td>
<td>2015</td>
</tr>
</tbody>
</table>
(Deep) Reinforcement Learning for Dialogue

User input (o) → Language understanding → Dialogue Manager $a = \pi(s)$ → Collect rewards $(s, a, r, s')$

$s \rightarrow$ Language (response) generation $\rightarrow a \rightarrow$ Optimize $Q(s, a)$

<table>
<thead>
<tr>
<th>Application</th>
<th>State</th>
<th>Action</th>
<th>Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task Completion Bots</td>
<td>User input + Context</td>
<td>Dialog act + slot_value</td>
<td>Task success rate</td>
</tr>
<tr>
<td>(Movies, Restaurants, ...)</td>
<td></td>
<td></td>
<td># of turns</td>
</tr>
<tr>
<td>Info Bots</td>
<td>Question + Context</td>
<td>Clarification questions, Answers</td>
<td>Relevance of answer</td>
</tr>
<tr>
<td>(Q&amp;A bot over KB, Web etc.)</td>
<td></td>
<td></td>
<td># of turns</td>
</tr>
<tr>
<td>Social Bot</td>
<td>Conversation history</td>
<td>Response</td>
<td>Engagement(?)</td>
</tr>
<tr>
<td>(XiaoIce)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A user simulator for RL and evaluation

- Robustness: automatic action selection based on uncertainty by RL
- Flexibility: allow user-initiated behaviors
- Reproducibility: a R&D setting that allows consistent comparisons of competing methods

[Li+ 17] https://github.com/MiuLab/TC-Bot
InfoBot as an interactive search engine

• Problem setting
  • User is looking for a piece of information from one or more tables/KBs
  • System must iteratively ask for user constraints (“slots”) to retrieve the answer

• A general rule-based approach
  • Given current beliefs, ask for slot with maximum uncertainty
  • Works well in most cases but,
    • Has no notion of what the user is likely to be looking for or likely to know
    • No principled way to deal with errors/uncertainty in language understanding
InfoBot as an interactive search engine

User simulator

Agent

User Utterance

System Response

Natural Language Understanding (NLU)

Acts/Entities

State Tracker/Belief Tracker

Query

Results

Database

Natural Language Generator (NLG)

Dialog Act

Dialog Policy

Dialog State
Deep Reinforcement Learning

User simulator

Agent

NLU

State Tracker

NLG

Dialog Policy

User Utterance

Reward

System Response

Acts/Entities

Dialog State

Backprop

Dialog Act

Query

Results

Database

Not Differentiable!
End-to-End Learning [Dhuwan+ 17]
Dual Exploration

• Agent should explore actions as well as KB outputs
  • Share similarities with RL Neural Turing Machines (NTU)
• Optimizing expected return

\[ J(\theta) = E_{a \sim \pi, I \sim p_{\tau}} \left[ \sum_{h=0}^{H} \gamma^h r_h \right] \]

• via REINFORCE

\[ \nabla_\theta J(\theta) = E_{a \sim \pi, I \sim p_{\tau}} \left[ \left( \nabla_\theta \log p(I) + \sum_{h=0}^{H} \nabla_\theta \log \pi(a_h) \right) \sum_{k=0}^{H} \gamma^k r_k \right] \]
Result on IMDB using KB-InfoBot w/ simulated users

<table>
<thead>
<tr>
<th>Agent</th>
<th>Success Rate</th>
<th>Avg Turns</th>
<th>Avg Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule-Soft</td>
<td>0.76</td>
<td>3.94</td>
<td>0.83</td>
</tr>
<tr>
<td>RL-Hard</td>
<td>0.75</td>
<td>3.07</td>
<td>0.86</td>
</tr>
<tr>
<td>RL-Soft</td>
<td>0.80</td>
<td>3.37</td>
<td>0.98</td>
</tr>
<tr>
<td>E2E-RL</td>
<td><strong>0.83</strong></td>
<td>3.27</td>
<td><strong>1.10</strong></td>
</tr>
</tbody>
</table>

Knowledge Base (head, relation, tail)
- (Groundhog Day, actor, Bill Murray)
- (Groundhog Day, release year, 1993)
- (Australia, actor, Nicole Kidman)
- (Mad Max: Fury Road, release year, 2015)
Results on real users
Composite task completion bot with Hierarchical RL

[17]

Travel Assistant

- Book Flight
- Reserve Restaurant
- Book Hotel

“subtasks”

Naturally solved by hierarchical RL
A hierarchical policy learner

Similar to HAM [Parr & Russell 98] and hierarchical DQN [Kulkarni+ 16]
Results on simulated and real users
Subgoal discovery for HRL:

Figure 3: Subgoals for the landmarks problem (Sutton et al., 1999). Though the solution with subgoals may not be optimal, having the subgoals could usually reduce the search space, and potentially accelerate the learning efficiency.

divided and conquer
The 4-room game

Figure 7: Termination probability visualization for the 4-room experiment. Each time the agent travels from the upper-left corner cell to the lower-right corner cell. The visualization shows the termination probabilities of the RNN generative models in the HRL training after the sequence segmentation process. Darker colors mean higher probabilities.
Summary

• The transition of NLP to neural approaches
• Neural approaches to MRC and QA
  • Knowledge representation and search in neural space
  • A case study: ReasoNet w/ long-term memory
  • Ongoing research: visualize the reasoning process in neural space
  • Learn more at Deep Learning for Machine Reading Comprehension
• An intelligent, human-like, open-domain conversational system
  • Dialogue as RL
  • Case study 1: InfoBot with end-to-end learning RL
  • Case study 2: Composite task completion bot with Hierarchical RL
  • Ongoing research: subgoal discovery for hierarchical RL
  • Learn more at Deep RL for goal-oriented dialogues
Contact Information:
www.microsoft.com/en-us/research/people/jfgao/

Collaborators:
Faisal Ahmed, Chris Brockett, Asli Celikyilmaz, Ming-Wei Chang, Weizhu Chen, Yun-Nung Chen, Li Deng, Bhuwan Dhingra, Bill Dolan, Michel Galley, Marjan Ghazvininejad, Xiaodong He, Po-Sen Huang, Sungjin Lee, Jiwei Li, Lihong Li, Xiujun Li, Zachary Lipton, Xiaodong Liu, Rangan Majumder, Nasrin Mostafazadeh, Baolin Peng, Mir Rosenberg, Yelong Shen, Alessandro Sordoni, Saurabh Tiwary, Lucy Vanderwende, Luan Yi, Scott Yih et al.