Semantic Parsing to Probabilistic Programs for Situated Question Answering

Jayant Krishnamurthy

Oyvind Tafjord  Aniruddha Kembhavi  Kenton Murray  Jonghyun Choi  Joel Grus
Situated Question Answering

Which animal is both predator and prey?

(A) hawk
(B) mouse
(C) cricket
(D) owl
Food Web Questions

How many organisms consume the mouse?
(A) 1  (B) 2  (C) 3  (D) 4

Which animal is both predator and prey?
(A) hawk  (B) mouse  (C) cricket  (D) owl

If all the rabbits die off, what will happen to the mountain lion population?
(A) it will increase  (B) it will decrease  (C) it will remain the same
Examples of Situated Question Answering

Visual QA
[Antol et al., 2015]

Q: what is her mustache made of?
A: bananas

Robot Commands
[Kollar et al., 2010]

Q: go down the hallway then right
A: (path through environment)

Text QA
[Berant et al., 2014]

Q: What can the splitting of water lead to?
A: Transfer of ions
Challenges

1. Noisy computer vision

2. Compositionality in language
   “predator and prey”

3. Interactions between question and diagram interpretations

4. Background knowledge
   Secondary consumer = an animal that eats animals that eat plants
Prior Work

Question: How many animals consume the mouse?

Semantic Parsing

Logical Form: 
(count (lambda (x) (eats x mouse))

Database:
- eats
  - weasel
  - mouse
- organism
  - mouse
  - shrubs

Execution: Answer 3

Answers:
- [Matuszek et al., 2012]
- [Krishnamurthy and Kollar, 2013]
- [Malinowski and Fritz, 2014]
Prior Work

Question: How many animals consume the mouse?

Environment

Encode

[Antol et al., 2015]
[Malinowski et al., 2015]
[Yang et al., 2016]
[Fukui et al., 2016]
[Andreas et al., 2016]
...

Encode

Combine

Softmax

Answer: 3
Outline

● Introduction

● Parsing to Probabilistic Programs [EMNLP 2016]
  ● Model
  ● Experiments

● Probabilistic Neural Programs [NAMPI 2016]
Question: How many animals consume the mouse?

Environment:
Semantic Parsing

Question: How many animals consume the mouse?

Semantic Parsing

Logical Form:
\[
\text{(count (lambda (x) (eats x mouse)))}
\]

[Zelle and Mooney, 1996] [Zettlemoyer and Collins 2005]
[Liang et al., 2011] [Kwiatkowski et al., 2013] [Artzi et al., 2013]
...
How many animals consume the mouse?

(count (lambda (x) (eats x mouse)))

2

Semantic Parsing

Logical Form

Database
eats
weasel
... mouse
... shrubs
... organism
mouse
weasel
... shrubs
...

[Zelle and Mooney, 1996] [Zettlemoyer and Collins 2005] [Liang et al., 2011] [Kwiatkowski et al., 2013] [Artzi et al., 2013] ...
Food Web Domain Theory

- Roles:
  organism, animal, plant, sun, decomposer, predator, prey, herbivore, carnivore, omnivore, primary-consumer, secondary-consumer, tertiary-consumer, top-predator

- Change Directions:
  increase, decrease, unchanged, affect

- Relations:
  eats, cause, has-role

- Miscellaneous:
  count, and, or, not, exists
If all the **frogs** died the population of **raccoon** would _?_

NP : S\NP :
frog \(\lambda x.\text{decrease}(x)\)

NP : (S\S)\NP :
raccoon \(\lambda x.\lambda e.\lambda f.\text{cause}(e, f(x))\)

S : \text{decrease(frog)}

(S\S) : \lambda e.\lambda f.\text{cause}(e, f(\text{raccoon}))

S : \lambda f.\text{cause}(<\text{decrease(frog)}, f(\text{raccoon}))

---

[Zelle and Mooney, 1996] [Zettlemoyer and Collins 2005]  
[Liang et al., 2011] [Kwiatkowski et al., 2013] [Artzi et al., 2013]  
[Krishnamurthy 2016]
Semantic Parsing

**Question**: How many animals consume the mouse?

**Logical Form**: \((\text{count} \ (\lambda (x) \ (\text{eats} \ x \ \text{mouse})))\)

**Execution**: 2

**Database**

- **eats**
  - mouse
  - shrubs

- **organism**
  - mouse
  - weasel
  - shrubs

References:

- Zelle and Mooney, 1996
- Zettlemoyer and Collins, 2005
- Liang et al., 2011
- Kwiatkowski et al., 2013
- Artzi et al., 2013

...
Question: How many animals consume the mouse?

Environment:

Semantic Parsing:

Logical Form:

(count (lambda (x) (eats x mouse)))
How many animals consume the mouse?

(count (lambda (x) (eats x mouse)))

Answer 3
;;; A normal program
(+ 1 1)

→ 2

[McCarthy, 1963]
[Goodman et al., 2008]
[Goodman and Stuhlmuller, 2014]
;;; choose represents a nondeterministic choice

(choose 1 2)
;;; choices can be nested
(+ (choose 1 2) (choose 3 4))
Probabilistic Programming

;; choices can be nested
(+ (choose 1 2) (choose 3 4))

[McCarthy, 1963]
[Goodman et al., 2008]
[Goodman and Stuhlmuller, 2014]
(define (eats x y)  
  ; first-cut approximation  
  (choose true false)  
)

; Helper functions
(define (plant x)  
  (eat x sun))

(define (count f)  
  (length (filter entities f)))

(define entities  
  (list weasel shrubs mouse ...))

Food web domain theory

Possible logical form executions
(define (eats x y)
  ; first-cut approximation
  (choose true false))

; Helper functions
(define (plant x)
  (eat x sun))

(define (count f)
  (length (filter entities f)))

(define entities
  (list weasel shrubs mouse ...))

Food web domain theory

Possible logical form executions
(define (eats x y)
  ; first-cut approximation
  (choose true false))

; Helper functions
(define (plant x)
  (eat x sun))

(define (count f)
  (length (filter entities f)))

(define entities
  (list weasel shrubs mouse ...))

Food web domain theory

Possible logical form executions
Question: How many animals consume the mouse?

Semantic Parsing:

Environment

Logical Form: 
\[
\text{count} \ (\lambda x \ (\text{eats} \ x \ \text{mouse}))
\]

Probabilistic Program Execution

Answer: 3
P³ as Loglinear Model

**Question**
How many animals consume the mouse?

**Logical Form**
(count (lambda (x) (eats x mouse)))

**Answer**
3

**Environment**

**Execution**

\[
P(\gamma|v, q; \theta) = \sum_{e, \ell, t} P(e, \ell, t|v, q; \theta)1(\text{ret}(e) = \gamma)
\]
P³ as Loglinear Model

\[ P(e, \ell, t|v, q; \theta) = \frac{1}{Z_{q,v}} f_{\text{ex}}(e, \ell|v; \theta_{\text{ex}}) f_{\text{prs}}(\ell, t|q; \theta_{\text{prs}}) \]
\[ f(\ell, t|s; \theta) = \exp\{\theta^T \phi(\ell, t, s)\} \]

\(\lambda f.\text{cause}(\text{decrease}(\text{frog}), f(\text{raccoon}))\)

"if all the frogs died ..."
Loglinear Execution Model

\[ f(e, \ell|d; \theta) = \exp\{\theta^T \phi(e, \ell, d)\} \]

\[ \lambda f. \text{cause}(\text{decrease}(frogs), f(raccoon)) \]
Execution Model
Execution Model

\[ 1 \exp\{\theta^T\} \]
Execution Model

\[ 1 \exp\{\theta^T\} \]

\[ \phi(\text{Weasel}) + \phi(\text{Mouse}) \]
Execution Model

\[ \exp\{\theta^T\} \]

\[ \phi(\text{Mouse}) + \phi(\text{Squirrel}) \]

(count (lambda (x) (eats x mouse)))

\[ \rightarrow 1 \]
How many animals consume the mouse?

(count (lambda (x) (eats x mouse)))

Answer 3
"if all the frogs died what will happen to the raccoons?"

\[
\text{Parser Beam Search}
\]

\[
(\lambda x \ (\text{eats} \ x \ \text{raccoon}))
\]

\[
(\lambda x \ (\text{cause} \ (\text{increase} \ x) \ \text{decrease} \ \text{raccoon}))
\]

\[
(\lambda x \ (\text{plant} \ x))
\]

\[
\ldots
\]

\[
\text{Execution Beam Search}
\]

\[
[\text{Goodman and Stuhlmuller, 2014}]
\]

\[
\rightarrow \ \{\text{frog}\}
\]

\[
\rightarrow \ \{\}
\]

\[
\rightarrow \ \{\text{frog, snake}\}
\]

\[
\rightarrow \ \{\text{rabbit}\}
\]

\[
\rightarrow \ \ldots
\]
Training

Question: "if all the frogs died ..."

Environment: $q^i$

Supervision Oracle: $c^i(e) = \begin{cases} 1 & \text{if execution } e \text{ is correct} \\ 0 & \text{otherwise} \end{cases}$

Answer: (A) increase (B) decrease (C) remain the same

Optimize Loglikelihood:

$$O(\theta) = \sum_{i=1}^{n} \log \sum_{e,l,t} c^i(e) P(e, l, t | q^i, v^i; \theta)$$

Labeled Environment:

<table>
<thead>
<tr>
<th>eats</th>
<th>organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>weasel mouse</td>
<td>mouse weasel</td>
</tr>
<tr>
<td>mouse shrubs</td>
<td>mouse shrubs</td>
</tr>
</tbody>
</table>

...
Outline

● Introduction

● Parsing to Probabilistic Programs  [EMNLP 2016]
  ● Model
  ● Experiments

● Probabilistic Neural Programs  [NAMPI 2016]
How will it most likely affect the ecosystem if the human population increases?
(A) sharks will increase
(B) fish take over
(C) everything will decrease (and be destroyed..)
(D) seagulls will eat sharks

Which of the organisms has more prey?
(A) stickleback fry
(B) epiphytic diatoms
(C) roach fry
(D) steelhead

What is a producer?
(A) fox
(B) hen
(C) plant
(D) python

5000 questions, 500 diagrams, 3/1/1 train/validation/test split
FoodWebs Accuracy

Accuracy

LSTM

60.3
FoodWebs Accuracy

Accuracy

[LSTM: 60.3] [DQA-NET: 59.3] [VQA: 56.5]

[Kembhavi et al., 2016]
FoodWebs Accuracy

Accuracy

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>LSTM</td>
<td>60.3</td>
</tr>
<tr>
<td>DQA-NET</td>
<td>59.3</td>
</tr>
<tr>
<td>VQA</td>
<td>56.5</td>
</tr>
<tr>
<td>Possible worlds</td>
<td>63.6</td>
</tr>
</tbody>
</table>

[Kembhavi et al., 2016]
FoodWebs Accuracy

![Graph showing accuracy of different models](image)

Accuracy

- LSTM: 60.3
- DQA-NET: 59.3
- VQA: 56.5
- Possible worlds: 63.6
- P3: 69.1

[Kembhavi et al., 2016]
FoodWebs Accuracy with Unseen Organisms

- LSTM: 34.7
- DQA-NET: 33
- VQA: 36.8
- Possible worlds: 50.8
- P3: 57.7
A blue colored coffee mug is placed very near to the computer on the table.

15 images, 284 natural language descriptions

[Krishnamurthy and Kollar, 2013]
Scene Experimental Results

Accuracy

<table>
<thead>
<tr>
<th>LSP</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>67</td>
<td>68</td>
</tr>
</tbody>
</table>

[Krishnamurthy and Kollar, 2013]
Scene Experimental Results

<table>
<thead>
<tr>
<th></th>
<th>LSP</th>
<th>P3</th>
<th>LSP</th>
<th>P3</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ Environments</td>
<td>67</td>
<td>68</td>
<td>70</td>
<td>75</td>
</tr>
<tr>
<td>+ Logical Forms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[ Krishnamurthy and Kollar, 2013 ]
Visual QA
[Antol et al., 2015]

Q: what is her mustache made of?
A: bananas

Chart Questions

Q: By how much did model X improve accuracy over the best baseline?
A: 5.5 points
Outline

- Introduction
- Parsing to Probabilistic Programs [EMNLP 2016]
  - Model
  - Experiments
- Probabilistic Neural Programs [NAMPI 2016]
(count (lambda (x)
    (eats x mouse)))

\[ \exp\{\theta^T\} \]
(count (lambda (x)
  (eats x mouse)))
\[ W = \text{parameter}(pW) \]
\[ b = \text{parameter}(pb) \]
\[ V = \text{parameter}(pV) \]
\[ a = \text{parameter}(pa) \]

\begin{align*}
x & = \text{vecInput}(2) \\
y & = \text{scalarInput}(0) \\
h & = \text{tanh}(W \times x + b) \\
y_{\text{pred}} & = \text{logistic}(V \times h + a)
\end{align*}
Computation Graphs for Structured Prediction

-Root-  This  time  around  ,  they 're moving  even  faster .
Computation Graphs for Structured Prediction

- Globally Normalized Models
- Cost-Sensitive Training
  - e.g., dynamic oracles for dependency parsing
- Reinforcement Learning

Not “free” with computation graphs:
Probabilistic Neural Programs

State Space of Discrete Choices

Computation Graphs for Neural Networks
val flip: Pp[Boolean] = choose(Seq(true, false), Seq(0.75, 0.25))
def mlp(featureVector: Tensor): Pp[CompGraphNode] = for {
    params <- param("params")
    bias <- param("bias")
    hidden = ((params * featureVector) + bias).tanh
    params2 <- param("params2")
    bias2 <- param("bias2")
    dist = (params2 * hidden) + bias2
} yield {
    dist
}
val nnFlip = for {
    dist <- mlp(featureVector)
    output <- choose(Array(false, true), dist)
} yield {
    output
}
val nnFlip = for {
  dist <- mlp(featureVector)
  flip <- choose(Array(false, true), dist)
  output <- if (flip) { value(1) }
  else { choose(Array(2, 3)) }
}

yield {
  output
}
val answerPnp: Pp[String] = answerQ("A thermometer …")
val answerDist = answerPnp.beamSearch(100, nnParams)

answerDist
temperature  0.7
fever         0.2
...

val data = List( (answerQ("The thermometer ..."), "temperature"),
               (answerQ("What season occurs when ..."), "summer"),
               ...
             )

val trainer = LoglikelihoodTrainer(...)

val trainedParams = trainer.train(data, initialParams, ...)

- Maximum likelihood w/ approximate inference
- Learning to search
- Reinforcement learning
Experiments

What happens to the snake population if the field mouse population decreases?
(A) it will increase  (B) it will decrease  (C) it will remain the same
Experiments

\[ \lambda f. \text{cause}(\text{decrease(mouse)}, f(\text{snake})) \]
Experiments
Experiments

- Loglinear: 8.6
- Multilayer Perceptron: 12.5
- Maxpool: 14.9
Conclusion

Which of the organisms has more prey?
(A) stickleback fry
(B) epiphytic diatoms
(C) roach fry
(D) steelhead

Code and data
http://github.com/allenai/pnp/