Lecture 16

Natural Language Programming

1. Machine Learning + Compiler
   • Specification $\rightarrow$ Agent
   • Training Data Synthesis

2. Algorithms and Results
Voice: Next-Generation Computer Interface

Queries of Knowledge

Dialogues: Intuitive software interfaces

End-User Task Automation
NLU: Manual Annotations

Too expensive!
Too many possible dialogues!
Must handle new sentences!

NLU: Pretrained Language Models

We need to ground them!
High-Level Compilers
From Schema to Agent
Domain Schema + Representative Data

restaurant {
  cuisine : String,
  rating : Number,
  ...
}

Natural Language

Training Data

Genie Tools

Pretrained Language Models

Semantic Parser

ThingTalk

[Campagna, PLDI 2019]
Scale

- Enable reuse
  - Standardized representation: ThingTalk
  - Open-source tools, training data, models

Empower 20M+ voice interface developers
Dialogues as Interfaces

“I like a Spanish restaurant in Palo Alto”
@yelp.restaurant(), geo == new Location("palo alto") && contains(categories, "spanish")

“I like a Spanish restaurant in Palo Alto”
@dialogue.execute;
@yelp.restaurant(), geo == new Location("palo alto") && contains(categories, "spanish")
#{results=[{id = "…", name = "La Bodeguita", price = moderate, ...}]}

“Please book a table for two”.
@yelp.make_reservation(book_people=2);

End-User Task Automation

“When I leave the house, turn off the lights.”
monitor(@get_gps()), location != $home
=> @light-bulb.set_power(power=off);
ich fand Dicke Wirtin in Berlin.
Möchten Sie eine Reservierung?

Bana yakınlarda güzel restoranlar bulun
Aklınızda belirli bir mutfak var mı?

"Köfte"

Ich möchte ein deutsches 5-
Sterne-Restaurant
Ja bitte für 8 Personen

Turkish

"Find me some good restaurants nearby"
"Do you have a specific cuisine in mind?"

"Burgers"

"Yes please, for 8 people"

"Ja bitte für 8 Personen"
Outline

1. Machine Learning + Compiler
   • Specification → Agent
   • Training Data Synthesis

2. Algorithms and Results
## Domain-Specific Paraphrases

- Naive paraphrases
  - Paraphrase whole sentences with a neural model (e.g. Seq2Seq)
  - Trained on a general-purpose paraphrasing dataset

<table>
<thead>
<tr>
<th>Synthetic</th>
<th>Paraphrased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search some cafeteria that have greater star than 3, and do not have smoking.</td>
<td>Search for a restaurant that has more than 3 stars and doesn't smoke.</td>
</tr>
<tr>
<td>Find restaurants close to my home.</td>
<td>Find restaurants near me.</td>
</tr>
<tr>
<td>Search for people who are employed by Stanford.</td>
<td></td>
</tr>
<tr>
<td>[greedy]</td>
<td>Look for people employed at Stanford.</td>
</tr>
<tr>
<td>[temperature=0.3]</td>
<td>Look for people who work at Stanford.</td>
</tr>
<tr>
<td>[temperature=1.2]</td>
<td>Find people at Stanford.</td>
</tr>
<tr>
<td>[temperature=1.5]</td>
<td>Actually, look for those who are currently employed at Stanford.</td>
</tr>
</tbody>
</table>
Filtering Noisy Paraphrases

1. Train a parser with the synthetic dataset
2. Generate noisy paraphrases of the synthetic dataset
3. Parse the paraphrases
4. Remove all paraphrases whose parses do not match the label
5. Add filtered paraphrases to the training set and train a new parser
6. Repeat

AutoQA: From Databases To Q&A Semantic Parsers With Only Synthetic Training Data
Silei Xu*, Sina J. Semnani*, Giovanni Campagna, Monica S. Lam
Neural Semantic Parser Model

- Pre-trained BERT encoder
- LSTM decoder

Schema2QA: High-Quality and Low-Cost Q&A Agents for the Structured Web
Silei Xu, Giovanni Campagna, Jian Li, and Monica S. Lam
In Proceedings of the 29th ACM International Conference on Information and Knowledge Management, October 2020
Question-Answering Results

Examples of Long-Tail Questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Alexa</th>
<th>Google</th>
<th>Siri</th>
<th>Genie</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show me restaurants rated at least 4 stars with at least 100 reviews</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Show restaurants in San Francisco rated higher than 4.5</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>What is the highest rated Chinese restaurant near Stanford?</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>How far is the closest 4 star restaurant?</td>
<td></td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Who works for W3C and went to Oxford?</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who worked for Google and lives in Palo Alto?</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Who graduated from Stanford and won a Nobel prize?</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>Who worked for at least 3 companies?</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Show me hotels with checkout time later than 12PM</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td>✔️</td>
</tr>
<tr>
<td>Which hotel has a pool in this area?</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Long-tail Restaurant Questions
Multi-Lingual Assistants

- Do we need to create generic templates for every language?
- Leverage machine translators
- Generate training data with local entities
- Using a novel neural aligner

<table>
<thead>
<tr>
<th>Lang.</th>
<th>Restaurant Queries with Localized Entities</th>
</tr>
</thead>
<tbody>
<tr>
<td>🇺🇸</td>
<td>look for 5 star restaurants that serve burgers</td>
</tr>
<tr>
<td>🇮🇩</td>
<td>cari 5 bintang restoran yang menyediakan burger</td>
</tr>
<tr>
<td>🇩🇪</td>
<td>suchen sie nach 5 sterne restaurants, die maultaschen servieren</td>
</tr>
<tr>
<td>🇪🇸</td>
<td>busque restaurantes de 5 estrellas que sirvan paella valenciana</td>
</tr>
<tr>
<td>🇯🇴</td>
<td>5 별의 레스토랑을 찾아보세요</td>
</tr>
<tr>
<td>🇫🇮</td>
<td>etsi 5 tähden ravintolaita, joissa tarjoillaan karjalanpiirakkaa</td>
</tr>
<tr>
<td>🇮🇹</td>
<td>cerca ristoranti a 5 stelle che servono bruschette</td>
</tr>
<tr>
<td>🇨と思っている 5つ星のレストランを調べてください</td>
<td></td>
</tr>
<tr>
<td>🇺🇸</td>
<td>poszukaj 5 gwiazdkowych restauracji, które serwują kotlet</td>
</tr>
<tr>
<td>🇹🇷</td>
<td>köfte servis eden 5 yıldızlı restoranları arayın</td>
</tr>
<tr>
<td>🇨🇳</td>
<td>搜索卖北京烤鸭的5星级餐厅</td>
</tr>
</tbody>
</table>
Results on Schema2QA Restaurants

Test with human-translated test data

Genie
- Few-shot training: synthetic + a few manually translated sentences

SOTA
- Translate foreign sentences to English with parameter substitution
- Use English semantic parser

Multi-lingual question answering with local entities: 1 DAY

Localizing Open-Ontology QA Semantic Parsers in a Day Using Machine Translation
Mehrad Moradshahi, Giovanni Campagna, Sina J. Semnani, Silei Xu, Monica S. Lam
Zooming In

- **User template**
- **Agent template**

**InfoQuestion**
What is the ___ of ___?

**SearchRequest**

**ProposeN**
I have ___ and ___. Both are ____.

**ProposeOne**
I have ___. It is a ____ ___ that ____.

**SearchRefine**
I don't like ___. Do you have something ___?

**AskAction**
I like that. Can you help me ___ it?

**SlotFillQuestion**

**ProvideInfo**
Combining with Domain Information

restaurant: price = "moderate", area = "centre"

ProposeOne
I have Pizza Express. It is a moderately priced restaurant that serves Italian food.

ProposeN
I have Pizza Express and Restaurant 17. Both are moderately priced.

AskAction
I like that. Can you help me book it?

SearchRefine
I don’t like Italian. Do you have something Indian?

InfoQuestion
What is the address of Pizza Express?

restaurant: price = “moderate”, area = “centre”, name = “Pizza Express”

restaurant: price = “moderate”, area = “centre”, food = “Indian”

restaurant: price = “moderate”, area = “centre”, address = “?”, name = “Pizza Express”
Contextual Neural Semantic Parser for Dialogues - Deployed in our Almond Assistant

User State
Exec: query Restaurant, food="Indian" & & ...
Best Dialogue Tracking Results

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Joint Accuracy (MultiWOZ 2.1)</strong></td>
<td></td>
</tr>
<tr>
<td>TRADE (Wu et al., 2019)</td>
<td>45.6</td>
</tr>
<tr>
<td>SUMBT (Lee et al., 2019a)</td>
<td>46.7</td>
</tr>
<tr>
<td>DSTQA (Zhou and Small, 2019)</td>
<td>51.2</td>
</tr>
<tr>
<td>DST-Picklist (Zhang et al., 2019a)</td>
<td>53.3</td>
</tr>
<tr>
<td>SST (Chen et al., 2020)</td>
<td>55.2</td>
</tr>
<tr>
<td>TripPy (Heck et al., 2020)</td>
<td>55.3</td>
</tr>
<tr>
<td>SimpleTOD (Hosseini-Asl et al., 2020)</td>
<td>55.7</td>
</tr>
<tr>
<td><strong>Turn-By-Turn Accuracy (Cleaned Test Set)</strong></td>
<td></td>
</tr>
<tr>
<td>Genie</td>
<td>85.3</td>
</tr>
</tbody>
</table>
RUSS
Rapid Universal Support Service

- Reads instructions → answer calls immediately
- Dataset: 80 customer services (741 instructions)
- 76.7% accuracy
- 69% of users prefer RUSS over following web instructions
- Trained with just synthetic data

[Nancy Xu et al. NAACL 21]
Summary: Tools for Developers

1. Per-field annotations
   - Attributes
   - Database + API Schemas

2. Sentence compositions
   - Pre-trained Model
   - Genie Grammar Templates

3. Domain-based paraphrases
   - Genie Dialogue Models

4. Dialogues
   - Neural Machine Translation

5. Multilingual data
General Lessons

- Control over conversational agents → not end-to-end neural networks
- Grounding pretrained neural networks to unleash its potential
  - Relationships between language
  - Common sense knowledge
- Computer-human natural language interface:
  → Semantic parser from natural language to code
- Compiler methodology supports reuse, empowers millions of users
  - Encapsulating knowhow in tools (Genie)
  - Standard formal language (Thingtalk)