What we’re about to discuss

- Introduction
- Task oriented dialogue systems
- Fully data-driven conversation models (chatbots)
- Conclusions and Research trends
- How to start?
Introduction:
What is a dialogue system?!?
Introduction: What is a dialogue system?!

- Emulating human conversation
- Answering questions on a wide range of topics
- Fulfilling complex tasks
Want to create a dialogue system?!!?
Want to create a dialogue system?!?
Want to create a dialogue system?!?

![Dialogue System Image]

**Chat Log:**

```
ELIZA: (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR?
AMIT: HI, I'M AMIT. WHAT DO YOU WANT TO TALK ABOUT?
ELIZA: CAN I TALK ABOUT MY PROBLEM?
AMIT: SURE.
ELIZA: DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA: PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT: THANKS FOR YOUR ADVICE.
ELIZA: NO MENTION.
AMIT: BYE.
ELIZA: BYE AND KEEP IN TOUCH...
```
Dialouge: What Kinds of Problems?

1. Question Answering
2. Task Completion
3. Social Chat
Which architecture do you choose?

Modular

Fully data-driven

input x

output y

Statistical model (e.g., neural)

E2E
What do you need to learn?!?

- NLP
- ML
- In particular DL
- RL can help
- IR
- Speech
Part two: Task-based dialogue system
Task-based dialogue system

An Overview

Why task-based dialogue system?
- We are tired!
- Conversation is a tool, not goal maybe.
- The faster the better!!

Example:
- Siri
- Alexa
- Cortana
- Google Now/Home
Task-based dialogue system

Slot filling Dialogues

- Simplest model
- Collect necessary information
- For each domain set of slot defined by experts
Task-based dialogue system

Dialogue Act

- **Dialogue Act Theory:** Interaction between agent and user
- **RL viewpoint**
  - Dialogue system is agent
  - Human is environment
  - Utterances are actions that can change state (api calls also can be)
- **Dialogue Act: Semantic between Natural Language and Action**
- Some dialogue acts may have slot as argument

<table>
<thead>
<tr>
<th>How many tickets do you need?</th>
<th>request(num_tickets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I want to watch it in Seattle</td>
<td>infrom(city=&quot;seattle&quot;)</td>
</tr>
<tr>
<td>What is the phone number?</td>
<td>request(phone_number)</td>
</tr>
</tbody>
</table>
Task-based dialogue system

Modular Architecture

<table>
<thead>
<tr>
<th>Natural Language Understanding</th>
<th>Converts the user utterances to dialogue act</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialogue Manager</td>
<td>Central controller, keep track of dialogue, select next action</td>
</tr>
<tr>
<td>Natural Language Generation</td>
<td>Convert the selected dialogue action to natural language</td>
</tr>
</tbody>
</table>
Task-based dialogue system

NLU

Three issues:

domain classification
intent determination
slot tagging

Example: *Show me morning flights from Boston to San Francisco on Tuesday*
Task-based dialogue system

Dialogue State Tracking

- It contains all information about what users is looking for at the current turn.
- Input for dialogue policy

User: I’m looking for a cheaper restaurant
inform(price=cheap)
System: Sure. What kind - and where?
User: Thai food, somewhere downtown
inform(price=cheap, food=Thai, area=centre)
System: The House serves cheap Thai food
User: Where is it?
inform(price=cheap, food=Thai, area=centre); request(address)
System: The House is at 106 Regent Street
Task-based dialogue system

Dialogue Policy

- Decide what action the system should take next
  - Dialogue act
  - Api call

\[ \hat{A}_i = \arg \max_{A_i \in A} P(A_i | (A_1, U_1, \ldots, A_{i-1}, U_{i-1}) \]

- Use RL
  - DQN whose entries correspond to all possible (dialogue-act, slot)

- More challenge:
  - Multi-domain dialogue
  - Composite-tasks
**Task-based dialogue system**

**NLG**

**Rule-based methods:**

Delexicalization

<table>
<thead>
<tr>
<th>recommend(restaurant name= Au Midi, neighborhood = midtown, cuisine = french)</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is a <strong>cuisine</strong> restaurant in <strong>neighborhood</strong> called <strong>restaurant_name</strong>.</td>
</tr>
<tr>
<td>There is a <strong>French</strong> restaurant in <strong>Midtown</strong> called <strong>Au Midi</strong>.</td>
</tr>
</tbody>
</table>

**Neural Methods:**
Task-based dialogue system

Evaluating

- We need to evaluate the whole of system
- Evaluation Metrics:
  - Task completion success
  - Time elapsed
- Simulation based evaluation:
  - Simulated user
    - Agenda based
    - Model Based
Task-based dialogue system

End to End learning

- Why modularity?
  - Flexibility to build module in independent way

- Why E2E?
  - Complex design of modular
  - Improve in individual module do not necessarily translate into improvement of whole system

- Supervised-Learning

- Reinforcement-Learning
  - Training data may expensive to obtain
  - exploration
Task-based dialogue system

Recent Frontier

- Beyond slot filling dialogues
  - Tutorial
  - Transaction
  - Negotiation
  - Multi-modal
- Weaker learning signal
  - Instead of having absolute judgement, one only requires a preferential input that indicates which one of two dialogues is better
- Related Areas
  - User simulation
  - Text games
  - ...
Fully data-driven conversation and chatbots
Fully data-driven conversation

introduction

- Trained entirely from data
- Without resorting to any expert knowledge
- Do not rely on four traditional components
- Easily scales to large free-from and open-domain datasets
- Conversation is a goal!
- Scenarios
  - Chitchat
  - Recommendation
  - ...

Fully data-driven conversation

Conversation Models

- Inspired by statistical machine translation
- Limitation: Bad look at data, contextually appropriate response
- RNN architectures
  - LSTM models
  - HRED models
  - Attention models
  - Pointer-Network models
    - Copy-net
Fully data-driven conversation
A Neural Conversation Model, 2015

Human: hello!
Machine: hello!
Human: how are you?
Machine: i’m good.
Human: what’s your name?
Machine: i’m julia.
Human: when were you born?
Machine: july 20th.
Human: what year were you born?
Human: where are you from?
Machine: i’m out in the boonies.
Human: are you a man or a woman?
Machine: i’m a woman.
Human: why are we here?
Machine: i’m not sure.
Human: okay, bye!
Machine: bye.
Human: see you later.
Machine: bye.
Fully data-driven conversation

A Neural Conversation Model, 2015

Human: who is skywalker ?
Machine: he is a hero .
Human: who is bill clinton ?
Machine: he 's a billionaire .
Human: is sky blue or black ?
Machine: blue .
Human: does a cat have a tail ?
Machine: yes .
Human: does a cat have a wing ?
Machine: no
Human: can a cat fly ?
Machine: no .
Human: how many legs does a cat have ?
Machine: four , i think .
Human: how many legs does a spider have ?
Machine: three , i think .
Human: how many legs does a centipede have ?
Machine: eight .
Human: what is the color of the sky ?
Machine: blue .
Human: what is the color of water ?
Machine: water .
Human: what is the color of blood ?
Machine: it is the same as a black eye .
Human: what is the usual color of a leaf ?
Machine: it is a green one .
Human: what is the color of a yellow car ?
Machine: yellow .
Human: how much is two plus two ?
Machine: four .

Human: what is the purpose of life ?
Machine: to serve the greater good .
Human: what is the purpose of living ?
Machine: to live forever .
Human: what is the purpose of existence ?
Machine: to find out what happens when we get to the planet earth .
Human: where are you now ?
Machine: i 'm in the middle of nowhere .
Human: what is the purpose of dying ?
Machine: to have a life .
Human: what is the purpose of being intelligent ?
Machine: to find out what it is .
Human: what is the purpose of emotions ?
Machine: i don 't know .
Fully data-driven conversation
A Neural Conversation Model, 2015

Problems begin!!!!

Conversation 7: Job and Personality

**Human:** what is your job?
**Machine:** i’m a lawyer.
**Human:** what do you do?
**Machine:** i’m a doctor.

**Human:** how much is ten minus two?
**Machine:** seventy-two.
Fully data-driven conversation

Challenges: Blandness

- Response Blandeness
  - Bland and uninformative
  - “I don’t know”!!! Or “Ok”

How to solve this problem?

\[
p(T|S) = \frac{P(S|T) \times P(T)}{P(S)}
\]

Hard to optimize! So only use in inference time

\[
\arg\max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}.
\]
Fully data-driven conversation

Challenges: Speaker-Consistency

<table>
<thead>
<tr>
<th>message</th>
<th>Where do you live now?</th>
</tr>
</thead>
<tbody>
<tr>
<td>response</td>
<td>I live in Los Angeles.</td>
</tr>
<tr>
<td>message</td>
<td>In which city do you live now?</td>
</tr>
<tr>
<td>response</td>
<td>I live in Madrid.</td>
</tr>
<tr>
<td>message</td>
<td>In which country do you live now?</td>
</tr>
<tr>
<td>response</td>
<td>England, you?</td>
</tr>
<tr>
<td>message</td>
<td>Where were you born?</td>
</tr>
<tr>
<td>response</td>
<td>I was born in Canada.</td>
</tr>
<tr>
<td>message</td>
<td>Where are you from?</td>
</tr>
<tr>
<td>response</td>
<td>England, you?</td>
</tr>
<tr>
<td>message</td>
<td>Where did you grow up?</td>
</tr>
<tr>
<td>response</td>
<td>I grew up in Texas.</td>
</tr>
<tr>
<td>message</td>
<td>How old are you?</td>
</tr>
<tr>
<td>response</td>
<td>16 and you?</td>
</tr>
<tr>
<td>message</td>
<td>What’s your age?</td>
</tr>
<tr>
<td>response</td>
<td>18.</td>
</tr>
<tr>
<td>message</td>
<td>What is your major?</td>
</tr>
<tr>
<td>response</td>
<td>I’m majoring in psychology</td>
</tr>
<tr>
<td>message</td>
<td>What did you study in college?</td>
</tr>
<tr>
<td>response</td>
<td>English lit.</td>
</tr>
</tbody>
</table>

Table 1: Inconsistent responses generated by a 4-layer SEQ2SEQ model trained on 25 million Twitter conversation snippets.
Fully data-driven conversation

Challenges: Speaker-Consistency
Fully data-driven conversation

Challenges: lack of grounding

- Most of chatbots are not grounded in real world!
- This limitation is also inherited from machine translation paradigm
- Some of ideas
  - Grounding system in the persona of speaker
  - Add textual knowledge source
  - Additional input drawn from the user environment, such as image
  - ...

Fully data-driven conversation

Evaluation

- Evaluation metrics:
  - Blue
  - ROUGE
  - METEOR
  - Delta Blue
- They are not appropriate for dialogue task
- Use GAN to evaluate
Fully data-driven conversation

Data

- Twitter
- Reddit
- OpenSubtitle
- Ubuntu
- Persona-Chat
- ...

Fully data-driven conversation
Open benchmarks

- Dialogue System Technology Challenge (DSTC)
- ConvAI Competition
- NTCIR STC
- Alexa Prize
- JD Dialogue Challenge
- ...
Conclusion and Research trends
Conclusion and Research Trends

<table>
<thead>
<tr>
<th>Question answering</th>
<th>SQUAD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WikiQA</td>
</tr>
<tr>
<td></td>
<td>CNN/DailyMail</td>
</tr>
<tr>
<td></td>
<td>COQA</td>
</tr>
<tr>
<td></td>
<td>NarrativeQA</td>
</tr>
<tr>
<td></td>
<td>Wikihop</td>
</tr>
<tr>
<td></td>
<td>Natural Questions</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task-based Dialogue Systems</th>
<th>Other genres</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Text games</td>
</tr>
<tr>
<td></td>
<td>E2E learning</td>
</tr>
<tr>
<td></td>
<td>User simulating</td>
</tr>
<tr>
<td></td>
<td>Application</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chatbots</th>
<th>New architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Response bladeness</td>
</tr>
<tr>
<td></td>
<td>Speaker-consistency</td>
</tr>
<tr>
<td></td>
<td>Knowledge grounded</td>
</tr>
<tr>
<td></td>
<td>Multi-model</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>
How To Start?!
How to start?!

- Learn deep learning and nlp!
- Learn Pytorch!!
- Read Papers about Dialogue
- Don’t Ignore Transfer Learning!
- Be Up to Date!
How to start?!
Read paper about dialogue

- **A Neural Conversational Model** *(ICML Deep Learning Workshop 2015)*
  Oriol Vinyals, Quoc Le

- **A Persona-Based Neural Conversation Model** *(ACL 2016)*
  Jiwei Li, Michel Galley, Chris Brockett, Georgios P. Spithourakis, Jianfeng Gao, Bill Dolan

- **A Simple, Fast Diverse Decoding Algorithm for Neural Generation** *(arXiv 2017)*
  Jiwei Li, Will Monroe, Dan Jurafsky

- **Neural Approaches to Conversational AI** *(arXiv 2018)*
  Jianfeng Gao, Michel Galley, Lihong Li

- **TransferTransfo: A Transfer Learning Approach for Neural Network Based Conversational Agents** *(NeurIPS 2018 CAI Workshop)*
  Thomas Wolf, Victor Sanh, Julien Chaumond, Clement Delangue

- **Wizard of Wikipedia: Knowledge-Powered Conversational agents** *(ICLR 2019)*

- **Learning to Speak and Act in a Fantasy Text Adventure Game** *(arXiv 2019)*
  Jack Urbanek, Angela Fan, Siddharth Karamcheti, Saachi Jain, Samuel Humeau, Emily Dinan, Tim Rocktäschel, Douwe Kiela, Arthur Szlam, Jason Weston

- **A knowledge-grounded neural conversation model**, Thirty-Second AAAI Conference on Artificial Intelligence
  Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, Michel Galley
How to start?!
Don’t Ignore Transfer Learning!

Question Answering on SQuAD1.1

Other methods vs. State-of-the-art methods
Transfer Learning

Elmo
Bert
OpenGPT
XLNET
...

(Image of Sesame Street characters)
How to start?!
Be Up to Date