How can I help you?
Outline

- **PART I.** Introduction & Background Knowledge
- **PART II.** Task-Oriented Dialogue Systems

................................. Break .................................

- **PART III.** Social Chat Bots
- **PART IV.** Evaluation
- **PART V.** Recent Trends and Challenges
Introduction
Outline

- PART I. Introduction & Background Knowledge
  - *Dialogue System Introduction*
  - Neural Network Basics
  - Reinforcement Learning
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Brief History of Dialogue Systems

Multi-modal systems
e.g., Microsoft MiPad, Pocket PC

TV Voice Search
e.g., Bing on Xbox

Virtual Personal Assistants

Early 1990s

Keyword Spotting
(e.g., AT&T)
System: “Please say collect, calling card, person, third number, or operator”

Early 2000s

Task-specific argument extraction
(e.g., Nuance, SpeechWorks)
User: “I want to fly from Boston to New York next week.”

2017

Intent Determination
(Nuance’s Emily™, AT&T HMIHY)
User: “Uh...we want to move...we want to change our phone line from this house to another house”

DARPA
CALO Project

Apple Siri
(2011)

Google Now (2012)
Google Assistant (2016)

Microsoft Cortana (2014)

Amazon Alexa/Echo (2014)

Facebook M & Bot (2015)

Google Home (2016)
Why We Need?

“I am smart”
“I have a question”
“I need to get this done”
“What should I do?”

Turing Test (“I” talk like a human)
Information consumption
Task completion
Decision support
Why We Need?

“I am smart”
“I have a question”
“I need to get this done”
“What should I do?”

- Turing Test (“I” talk like a human)
- Information consumption
- Task completion
- Decision support

- What is the employee review schedule?
- Which room is the dialogue tutorial in?
- When is the IJCNLP 2017 conference?
- What does NLP stand for?
Why We Need?

“I am smart”
“I have a question”
“I need to get this done”
“What should I do?”

Turing Test (“I” talk like a human)
Information consumption
Task completion
Decision support

• Book me the flight from Seattle to Taipei
• Reserve a table at Din Tai Fung for 5 people, 7PM tonight
• Schedule a meeting with Bill at 10:00 tomorrow.
Why We Need?

“I am smart”
“I have a question”
“I need to get this done”
“What should I do?”

Turing Test (“I” talk like a human)
Information consumption
Task completion
Decision support

• *Is this product worth to buy?*
Why We Need?

“I am smart”
“I have a question”
“I need to get this done”
“What should I do?”

Turing Test ("I" talk like a human)
Information consumption
Task completion
Decision support

Task-Oriented Dialogues

Material: http://opendialogue.miulab.tw
Language Empowering Intelligent Assistant

Apple Siri (2011)
Google Now (2012)
Facebook M & Bot (2015)
Google Home (2016)
Amazon Alexa/Echo (2014)
Microsoft Cortana (2014)
Google Assistant (2016)
Apple HomePod (2017)

Material: http://opendialogue.miulab.tw
Intelligent Assistants

Task-Oriented

Engaging (social bots)
Why Natural Language?

- Global Digital Statistics (2017 January)

- Total Population 7.48B
- Internet Users 3.77B
- Active Social Media Users 2.79B
- Unique Mobile Users 4.92B
- Active Mobile Social Users 2.55B

The more natural and convenient input of devices evolves towards speech.
Spoken Dialogue System (SDS)

- **Spoken dialogue systems** are intelligent agents that are able to help users finish tasks more efficiently via spoken interactions.
- **Spoken dialogue systems** are being incorporated into various devices (smart-phones, smart TVs, in-car navigating system, etc).

Good dialogue systems assist users to access information conveniently and finish tasks efficiently.
A bot is responsible for a “single” domain, similar to an app. Users can initiate dialogues instead of following the GUI design.
GUI v.s. CUI (Conversational UI)

https://github.com/enginebai/Movie-lol-android

Material: http://opendialogue.miulab.tw
# GUI v.s. CUI (Conversational UI)

<table>
<thead>
<tr>
<th></th>
<th>Website/APP’s GUI</th>
<th>Msg’s CUI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Situation</strong></td>
<td>Navigation, no specific goal</td>
<td>Searching, with specific goal</td>
</tr>
<tr>
<td><strong>Information Quantity</strong></td>
<td>More</td>
<td>Less</td>
</tr>
<tr>
<td><strong>Information Precision</strong></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td><strong>Display</strong></td>
<td>Structured</td>
<td>Non-structured</td>
</tr>
<tr>
<td><strong>Interface</strong></td>
<td>Graphics</td>
<td>Language</td>
</tr>
<tr>
<td><strong>Manipulation</strong></td>
<td>Click</td>
<td>mainly use texts or speech as input</td>
</tr>
<tr>
<td><strong>Learning</strong></td>
<td>Need time to learn and adapt</td>
<td>No need to learn</td>
</tr>
<tr>
<td><strong>Entrance</strong></td>
<td>App download</td>
<td>Incorporated in any msg-based interface</td>
</tr>
<tr>
<td><strong>Flexibility</strong></td>
<td>Low, like machine manipulation</td>
<td>High, like converse with a human</td>
</tr>
</tbody>
</table>

[Material: http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
Two Branches of Dialogue Systems

**Task-Oriented Bot**
- Personal assistant, helps users achieve a certain task
- Combination of rules and statistical components
  - POMDP for spoken dialog systems (Williams and Young, 2007)
  - End-to-end trainable task-oriented dialogue system (Wen et al., 2016)
  - End-to-end reinforcement learning dialogue system (Li et al., 2017; Zhao and Eskenazi, 2016)

**Chit-Chat Bot**
- No specific goal, focus on natural responses
- Using variants of seq2seq model
  - A neural conversation model (Vinyals and Le, 2015)
  - Reinforcement learning for dialogue generation (Li et al., 2016)
  - Conversational contextual cues for response ranking (Al-Rfou et al., 2016)
Task-Oriented Dialogue System (Young, 2000)

Speech Signal

Hypothesis
are there any action movies to see this weekend

Text Input
Are there any action movies to see this weekend?

Speech Recognition

Language Understanding (LU)
- Domain Identification
- User Intent Detection
- Slot Filling

Semantic Frame
request_movie
genre=action, date=this weekend

Natural Language Generation (NLG)

Text response
Where are you located?

System Action/Policy
request_location

Dialogue Management (DM)
- Dialogue State Tracking (DST)
- Dialogue Policy

Backend Action / Knowledge Providers

Material: http://opendialogue.miulab.tw
Outline

- PART I. Introduction & Background Knowledge
  - Dialogue System Introduction
  - *Neural Network Basics*
  - Reinforcement Learning
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Machine Learning ≈ Looking for a Function

- **Speech Recognition**
  \[ f(\text{speech}) = \text{“你好 (Hello)”} \]

- **Image Recognition**
  \[ f(\text{image}) = \text{cat} \]

- **Go Playing**
  \[ f(\text{game board}) = \text{5-5 (next move)} \]

- **Chat Bot**
  \[ f(\text{“Where is IJCNLP?”}) = \text{“The location is...”} \]

Given a large amount of data, the machine learns what the function \( f \) should be.
Deep learning is a type of machine learning approaches, called “neural networks”.
A Single Neuron

\[ y = h_{w,b}(x) = \sigma(w^T x + b) \]

Activation function

\[ \sigma(z) = \frac{1}{1 + e^{-z}} \]

Sigmoid function

\[ w, b \] are the parameters of this neuron
A Single Neuron

A single neuron can only handle binary classification

$$f : \mathbb{R}^N \rightarrow \mathbb{R}^M$$

$$y = \begin{cases} \text{"2"} & y \geq 0.5 \\ \text{not "2"} & y < 0.5 \end{cases}$$
A Layer of Neurons

- Handwriting digit classification

A layer of neurons can handle multiple possible output, and the result depends on the max one.

$$f : \mathbb{R}^N \rightarrow \mathbb{R}^M$$

10 neurons/10 classes

Which one is max?
Deep Neural Networks (DNN)

- Fully connected feedforward network

\[ f : \mathbb{R}^N \rightarrow \mathbb{R}^M \]

![Diagram of Deep Neural Networks](image)

**Input**\( x \)

**Layer 1**\( y_1 \)

**Layer 2**\( y_2 \)

**Layer L**\( y_M \)

**Output**\( y \)

**Deep NN: multiple hidden layers**
Recurrent Neural Network (RNN)

\[ s_t = \sigma(W s_{t-1} + U x_t) \quad \sigma(\cdot): \text{tanh, ReLU} \]

\[ o_t = \text{softmax}(V s_t) \]

RNN can learn accumulated sequential information (time-series)
Outline

- PART I. Introduction & Background Knowledge
  - Dialogue System Introduction
  - Neural Network Basics
  - Reinforcement Learning
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges

Material: http://opendialogue.miulab.tw
Reinforcement Learning

- RL is a general purpose framework for **decision making**
  - RL is for an *agent* with the capacity to *act*
  - Each *action* influences the agent’s future *state*
  - Success is measured by a scalar *reward* signal
  - Goal: *select actions to maximize future reward*
Scenario of Reinforcement Learning

Agent learns to take actions to maximize expected reward.

Observation $o_t$  |  Action $a_t$
---|---
Reward $r_t$

If win, reward = 1
If loss, reward = -1
Otherwise, reward = 0

Next Move

Material: [http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
Supervised v.s. Reinforcement

- **Supervised**
  - Learning from teacher
  - Say “Hello”
  - Say “Bye bye”

- **Reinforcement**
  - Learning from critics
  - Hello 😊
  - Say “Bad”

Material: [http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
Sequential Decision Making

- Goal: select actions to maximize total future reward
  - Actions may have long-term consequences
  - Reward may be delayed
  - It may be better to sacrifice immediate reward to gain more long-term reward

Material: [http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
Deep Reinforcement Learning

- Observation
- Function Input
- DNN
- Reward
- Environment
- Used to pick the best function
- Action
- Function Output

Material: http://opendialogue.miulab.tw
Reinforcing Learning

- Start from state $s_0$
- Choose action $a_0$
- Transit to $s_1 \sim P(s_0, a_0)$
- Continue...

$$s_0 \xrightarrow{a_0} s_1 \xrightarrow{a_1} s_2 \xrightarrow{a_2} s_3 \xrightarrow{a_3} \ldots$$

- Total reward:

$$R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots$$

**Goal**: select actions that maximize the expected total reward

$$\mathbb{E}[R(s_0) + \gamma R(s_1) + \gamma^2 R(s_2) + \cdots]$$
Reinforcement Learning Approach

- **Policy-based RL**
  - Search directly for optimal policy \( \pi^* \)

- **Value-based RL**
  - Estimate the optimal value function \( Q^*(s, a) \)

- **Model-based RL**
  - Build a model of the environment
  - Plan (e.g. by lookahead) using model

\( \pi^* \) is the policy achieving maximum future reward

\( Q^*(s, a) \) is maximum value achievable under any policy
Task-Oriented Dialogue Systems
Task-Oriented Dialogue System (Young, 2000)

Speech Signal

Hypothesis
are there any action movies to see this weekend

Text Input
Are there any action movies to see this weekend?

Speech Recognition

Language Understanding (LU)
- Domain Identification
- User Intent Detection
- Slot Filling

Semantic Frame
request_movie
genre=action, date=this weekend

System Action/Policy
request_location

Natural Language Generation (NLG)

Dialogue Management (DM)
- Dialogue State Tracking (DST)
- Dialogue Policy

Backend Action / Knowledge Providers

Text response
Where are you located?
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management – Dialogue State Tracking (DST)
  - Dialogue Management – Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Language Understanding (LU)

- Pipelined

1. Domain Classification
2. Intent Classification
3. Slot Filling
Mainly viewed as an utterance classification task

- Given a collection of utterances $u_i$ with labels $c_i$, $D=\{(u_1,c_1),\ldots,(u_n,c_n)\}$ where $c_i \in C$, train a model to estimate labels for new utterances $u_k$. 

find me a cheap taiwanese restaurant in oakland

Movies	Find_movie
Restaurants	Buy_tickets
Sports	Find_restaurant
Weather	Book_table
Music	Find_lyrics
...
...
DNN for Domain/Intent Classification (Ravuri & Stolcke, 2015)

- RNN and LSTMs for utterance classification

Intent decision after reading all words performs better
DNN for Dialogue Act Classification (Lee & Dernoncourt, 2016)

- RNN and CNNs for dialogue act classification
LU – Slot Filling

As a sequence tagging task

• Given a collection tagged word sequences,  \( S = \{ ((w_{1,1}, w_{1,2}, \ldots, w_{1,n1}), (t_{1,1}, t_{1,2}, \ldots, t_{1,n1})), ((w_{2,1}, w_{2,2}, \ldots, w_{2,n2}), (t_{2,1}, t_{2,2}, \ldots, t_{2,n2})) \ldots \} \)

where \( t_i \in M \), the goal is to estimate tags for a new word sequence.

flights from Boston to New York today

<table>
<thead>
<tr>
<th>flights</th>
<th>from</th>
<th>Boston</th>
<th>to</th>
<th>New</th>
<th>York</th>
<th>today</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>O</td>
<td>B-city</td>
<td>O</td>
<td>B-city</td>
<td>I-city</td>
<td>O</td>
</tr>
<tr>
<td>O</td>
<td>O</td>
<td>B-dept</td>
<td>O</td>
<td>B-arrival</td>
<td>I-arrival</td>
<td>B-date</td>
</tr>
</tbody>
</table>
RNN for Slot Tagging – I (Yao et al, 2013; Mesnil et al, 2015)

- Variations:
  - a. RNNs with LSTM cells
  - b. Input, sliding window of n-grams
  - c. Bi-directional LSTMs
RNN for Slot Tagging – II (Kurata et al., 2016; Simonnet et al., 2015)

- Encoder-decoder networks
  - Leverages sentence level information

- Attention-based encoder-decoder
  - Use of attention (as in MT) in the encoder-decoder network
  - Attention is estimated using a feed-forward network with input: $h_t$ and $s_t$ at time $t$
RNN for Slot Tagging – III (Jaech et al., 2016; Tafforeau et al., 2016)

- Multi-task learning
  - Goal: exploit data from domains/tasks with a lot of data to improve ones with less data
  - Lower layers are shared across domains/tasks
  - Output layer is specific to task
Joint Segmentation and Slot Tagging (Zhai et al., 2017)

- Encoder that segments
- Decoder that tags the segments
Joint Semantic Frame Parsing

- Slot filling and intent prediction in the same output sequence
  - Sequence-based (Hakkani-Tur+ 16)

- Intent prediction and slot filling are performed in two branches
  - Parallel-based (Liu+ 16)
Contextual LU

Domain Identification $\rightarrow$ Intent Prediction $\rightarrow$ Slot Filling

$D$ communication $\rightarrow$ $I$ send_email

$U$ just sent email to bob about fishing this weekend

$S$ B-contact_name B-subject I-subject I-subject

$\Rightarrow$ send_email(contact_name="bob", subject="fishing this weekend")

$U_1$ send email to bob

$S_1$ B-contact_name

$\Rightarrow$ send_email(contact_name="bob")

$U_2$ are we going to fish this weekend

$S_2$ B-message I-message I-message I-message I-message

$\Rightarrow$ send_email(message="are we going to fish this weekend")
Contextual LU

- User utterances are highly ambiguous in isolation

Restaurant Booking

Book a table for 10 people tonight.

Which restaurant would you like to book a table for?

Cascal, for 6.

#people time
Contextual LU  (Bhargava et al., 2013; Hori et al., 2015)

- Leveraging contexts
  - Used for individual tasks

- Seq2Seq model
  - Words are input one at a time, tags are output at the end of each utterance

- Extension: LSTM with speaker role dependent layers
End-to-End Memory Networks (Sukhbaatar et al, 2015)

U: “i d like to purchase tickets to see deepwater horizon”
S: “for which theatre”
U: “angelika”
S: “you want them for angelika theatre?”
U: “yes angelika”
S: “how many tickets would you like ?”
U: “3 tickets for saturday”
S: “What time would you like ?”
U: “Any time on saturday is fine”
S: “okay , there is 4:10 pm , 5:40 pm and 9:20 pm”
U: “Let’s do 5:40”
E2E MemNN for Contextual LU (Chen et al., 2016)

1. **Sentence Encoding**
   \[ m_i = \text{RNN}_{\text{mem}}(x_i) \]
   \[ u = \text{RNN}_{\text{in}}(c) \]

2. **Knowledge Attention**
   \[ p_i = \text{softmax}(u^T m_i) \]

3. **Knowledge Encoding**
   \[ h = \sum_i p_i m_i \]
   \[ o = W_{\text{kg}}(h + u) \]

**Idea:** additionally incorporating contextual knowledge during slot tagging → track dialogue states in a latent way
Analysis of Attention

U: “I'd like to purchase tickets to see Deepwater Horizon”
S: “for which theatre”
U: “Angelika”
S: “you want them for Angelika theatre?”
U: “yes Angelika”
S: “how many tickets would you like?”
U: “3 tickets for saturday”
S: “What time would you like?”
U: “Any time on saturday is fine”
S: “okay, there is 4:10 pm, 5:40 pm and 9:20 pm”
U: “Let’s do 5:40”
Sequential Dialogue Encoder Network (Bapna et al., 2017)

- Past and current turn encodings input to a feed forward network
**Structural LU** *(Chen et al., 2016)*

- **K-SAN:** prior knowledge as a teacher

![Diagram of Knowledge Encoding Module](image-url)
Structural LU (Chen et al., 2016)

- Sentence structural knowledge stored as memory

**Sentence**
show me the flights from seattle to san francisco

**Syntax (Dependency Tree)**
- ROOT
- show
- 1. me
- 2. the
- 3. seattle
- 4. san

**Semantics (AMR Graph)**
- show
- 1. you
- 2. city
- 3. Seattle
- 4. flight
- city
- San Francisco
Structural LU \cite{Chen2016}

- Sentence structural knowledge stored as memory

Using \textit{less training data} with K-SAN allows the model pay the similar attention to the \textit{salient substructures} that are important for tagging.
LU Importance (Li et al., 2017)

- Compare different types of LU errors

Sensitivity to Intent Error

Sensitivity to Slot Error

Slot filling is more important than intent detection in language understanding
LU Evaluation

- **Metrics**
  - Sub-sentence-level: intent accuracy, slot F1
  - Sentence-level: whole frame accuracy
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management – Dialogue State Tracking (DST)
  - Dialogue Management – Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - End-to-End Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Elements of Dialogue Management

(Figure from Gašić)

Dialogue State Tracking

- What the system says:
- What the user wants:
- What the system hears:

actions: a1, a2, a3, at-1
states: S1, S2, S3, ST-1, ST
observations: O1, O2, O3, OT-1, OT

dialogue turns

Material: http://opendialogue.miulab.tw
Dialogue State Tracking (DST)

- Maintain a probabilistic distribution instead of a 1-best prediction for better robustness to SLU errors or ambiguous input

<table>
<thead>
<tr>
<th>Slot</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td># people</td>
<td>5 (0.5)</td>
</tr>
<tr>
<td>time</td>
<td>5 (0.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Slot</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td># people</td>
<td>3 (0.8)</td>
</tr>
<tr>
<td>time</td>
<td>5 (0.8)</td>
</tr>
</tbody>
</table>
Multi-Domain Dialogue State Tracking (DST)

- A full representation of the system's belief of the user's goal at any point during the dialogue
- Used for making API calls

<table>
<thead>
<tr>
<th>Movies</th>
<th>Restaurants</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/15/16</td>
<td>Date</td>
</tr>
<tr>
<td>6 pm</td>
<td>11/15/16</td>
</tr>
<tr>
<td>7 pm</td>
<td></td>
</tr>
<tr>
<td>8 pm</td>
<td>Time</td>
</tr>
<tr>
<td>9 pm</td>
<td>6:30 pm</td>
</tr>
<tr>
<td>2</td>
<td>7 pm</td>
</tr>
<tr>
<td>3</td>
<td>7:30 pm</td>
</tr>
<tr>
<td>11/15/16</td>
<td>Cuisine</td>
</tr>
<tr>
<td></td>
<td>Mexican</td>
</tr>
<tr>
<td>Century</td>
<td>Restaurant</td>
</tr>
<tr>
<td>Trolls</td>
<td>Vive Sol</td>
</tr>
<tr>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>

Do you wanna take Angela to go see a movie tonight? 
Sure, I will be home by 6.
Let’s grab dinner before the movie.
How about some Mexican? 
Let’s go to Vive Sol and see Inferno after that.
Angela wants to watch the Trolls movie. Ok. Lets catch the 8 pm show.
## Dialog State Tracking Challenge (DSTC)


<table>
<thead>
<tr>
<th>Challenge</th>
<th>Type</th>
<th>Domain</th>
<th>Data Provider</th>
<th>Main Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSTC1</td>
<td>Human-Machine</td>
<td>Bus Route</td>
<td>CMU</td>
<td>Evaluation Metrics</td>
</tr>
<tr>
<td>DSTC2</td>
<td>Human-Machine</td>
<td>Restaurant</td>
<td>U. Cambridge</td>
<td>User Goal Changes</td>
</tr>
<tr>
<td>DSTC3</td>
<td>Human-Machine</td>
<td>Tourist Information</td>
<td>U. Cambridge</td>
<td>Domain Adaptation</td>
</tr>
<tr>
<td>DSTC4</td>
<td>Human-Human</td>
<td>Tourist Information</td>
<td>I2R</td>
<td>Human Conversation</td>
</tr>
<tr>
<td>DSTC5</td>
<td>Human-Human</td>
<td>Tourist Information</td>
<td>I2R</td>
<td>Language Adaptation</td>
</tr>
</tbody>
</table>
NN-Based DST (Henderson et al., 2013; Mrkšić et al., 2015; Mrkšić et al., 2016)
Neural Belief Tracker (Mrkšić et al., 2016)
DST Evaluation

- Dialogue State Tracking Challenges
  - DSTC2-3, human-machine
  - DSTC4-5, human-human

- Metric
  - Tracked state accuracy with respect to user goal
  - Recall/Precision/F-measure individual slots
Outline

PART I. Introduction & Background Knowledge

PART II. Task-Oriented Dialogue Systems
- Spoken/Natural Language Understanding (SLU/NLU)
- Dialogue Management – Dialogue State Tracking (DST)
- *Dialogue Management – Dialogue Policy Optimization*
- Natural Language Generation (NLG)
- End-to-End Task-Oriented Dialogue Systems

PART III. Social Chat Bots

PART IV. Evaluation

PART V. Recent Trends and Challenges
Elements of Dialogue Management

Dialogue Policy Optimization

(Figure from Gašić)
Dialogue Policy Optimization

- Dialogue management in a RL framework

Optimized dialogue policy selects the best action that can maximize the future reward. Correct rewards are a crucial factor in dialogue policy training.
Reward for RL \(\cong\) Evaluation for System

- Dialogue is a special RL task
  - Human involves in interaction and rating (evaluation) of a dialogue
  - Fully human-in-the-loop framework
- Rating: correctness, appropriateness, and adequacy

<table>
<thead>
<tr>
<th>Rating Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert rating</td>
<td>high quality, high cost</td>
</tr>
<tr>
<td>User rating</td>
<td>unreliable quality, medium cost</td>
</tr>
<tr>
<td>Objective rating</td>
<td>Check desired aspects, low cost</td>
</tr>
</tbody>
</table>
Reinforcement Learning for Dialogue Policy Optimization

User input (o) ➔ Language understanding ➔ Dialogue Policy \( a = \pi(s) \) ➔ Collect rewards \((s, a, r, s')\) ➔ Optimize \(Q(s, a)\)

Response

<table>
<thead>
<tr>
<th>Type of Bots</th>
<th>State</th>
<th>Action</th>
<th>Reward</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social ChatBots</td>
<td>Chat history</td>
<td>System Response</td>
<td># of turns maximized; Intrinsically motivated reward</td>
</tr>
<tr>
<td>InfoBots (interactive Q/A)</td>
<td>User current question + Context</td>
<td>Answers to current question</td>
<td>Relevance of answer; # of turns minimized</td>
</tr>
<tr>
<td>Task-Completion Bots</td>
<td>User current input + Context</td>
<td>System dialogue act w/ slot value (or API calls)</td>
<td>Task success rate; # of turns minimized</td>
</tr>
</tbody>
</table>

Goal: develop a generic deep RL algorithm to learn dialogue policy for all bot categories
Dialogue Reinforcement Learning Signal

- Typical reward function
  - -1 for per turn penalty
  - Large reward at completion if successful

- Typically requires domain knowledge
  - ✔ Simulated user
  - ✔ Paid users (Amazon Mechanical Turk)
  - ✖ Real users

The user simulator is usually required for dialogue system training before deployment
Neural Dialogue Manager \((\text{Li et al., 2017})\)

- Deep Q-network for training DM policy
  - Input: current semantic frame observation, database returned results
  - Output: system action

![Diagram of Neural Dialogue Manager](Material: http://deepdialogue.miulab.tw)
SL + RL for Sample Efficiency (Su et al., 2017)

- Issue about RL for DM
  - slow learning speed
  - cold start

- Solutions
  - Sample-efficient actor-critic
    - Off-policy learning with experience replay
    - Better gradient update
  - Utilizing supervised data
    - Pretrain the model with SL and then fine-tune with RL
    - Mix SL and RL data during RL learning
    - Combine both
Online Training (Su et al., 2015; Su et al., 2016)

- Policy learning from real users
  - Infer reward directly from dialogues (Su et al., 2015)
  - User rating (Su et al., 2016)

- Reward modeling on user binary success rating
Interactive RL for DM (Shah et al., 2016)

Use a third agent for providing interactive feedback to the DM
Dialogue Management Evaluation

- **Metrics**
  - Turn-level evaluation: system action accuracy
  - Dialogue-level evaluation: task success rate, reward
Outline

- PART I. Introduction & Background Knowledge
- **PART II. Task-Oriented Dialogue Systems**
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management – Dialogue State Tracking (DST)
  - Dialogue Management – Dialogue Policy Optimization
  - *Natural Language Generation (NLG)*
  - End-to-End Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Natural Language Generation (NLG)

- Mapping semantic frame into natural language
  
  \[
  \text{inform(name=Seven\_Days, foodtype=Chinese)}
  \]

  Seven Days is a nice Chinese restaurant
Template-Based NLG

- Define a set of rules to map frames to NL

<table>
<thead>
<tr>
<th>Semantic Frame</th>
<th>Natural Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>confirm()</td>
<td>“Please tell me more about the product you are looking for.”</td>
</tr>
<tr>
<td>confirm(area=$V)</td>
<td>“Do you want somewhere in the $V?”</td>
</tr>
<tr>
<td>confirm(food=$V)</td>
<td>“Do you want a $V restaurant?”</td>
</tr>
<tr>
<td>confirm(food=$V,area=$W)</td>
<td>“Do you want a $V restaurant in the $W.”</td>
</tr>
</tbody>
</table>

**Pros:** simple, error-free, easy to control

**Cons:** time-consuming, poor scalability
Plan-Based NLG (Walker et al., 2002)

- Divide the problem into pipeline

  - Sentence Plan Generator
  - Sentence Plan Reranker
  - Surface Realizer

```
Inform(
  name=Z_House,
  price=cheap
)
```

- **Statistical sentence plan generator** (Stent et al., 2009)
- **Statistical surface realizer** (Dethlefs et al., 2013; Cuayáhuitl et al., 2014; ...)

**Pros:** can model complex linguistic structures

**Cons:** heavily engineered, require domain knowledge
Class-Based LM NLG (Oh and Rudnicky, 2000)

- Class-based language modeling
  \[
  P(X \mid c) = \sum_{t} \log p(x_t \mid x_0, x_1, \cdots, x_{t-1}, c)
  \]

- NLG by decoding
  \[
  X^* = \arg \max_{X} P(X \mid c)
  \]

Pros: easy to implement/understand, simple rules

Cons: computationally inefficient
RNN-Based LM NLG (Wen et al., 2015)

Input: Inform(name=Din Tai Fung, food=Taiwanese)

dialogue act 1-hot representation

conditioned on the dialogue act

Output: delexicalisation

Slot weight tying

Slot weight tying
Handling Semantic Repetition

- Issue: semantic repetition
  - Din Tai Fung is a great Taiwanese restaurant that serves Taiwanese.
  - Din Tai Fung is a child friendly restaurant, and also allows kids.

- Deficiency in either model or decoding (or both)

- Mitigation
  - Post-processing rules (Oh & Rudnicky, 2000)
  - Gating mechanism (Wen et al., 2015)
  - Attention (Mei et al., 2016; Wen et al., 2015)
Semantic Conditioned LSTM  
(Wen et al., 2015)

Original LSTM cell

\[ i_t = \sigma(W_{wi}x_t + W_{hi}h_{t-1}) \]
\[ f_t = \sigma(W_{wf}x_t + W_{hf}h_{t-1}) \]
\[ o_t = \sigma(W_{wo}x_t + W_{ho}h_{t-1}) \]
\[ c_t = \tanh(W_{wc}x_t + W_{hc}h_{t-1}) \]
\[ h_t = o_t \odot \tanh(c_t) \]

Dialogue act (DA) cell

\[ r_t = \sigma(W_{wr}x_t + W_{hr}h_{t-1}) \]
\[ d_t = r_t \odot d_{t-1} \]

Modify \( C_t \)

\[ c_t = f_t \odot c_{t-1} + i_t \odot \hat{c}_t + \tanh(W_{dc}d_t) \]

Idea: using gate mechanism to control the generated semantics (dialogue act/slots)
Structural NLG (Dušek and Jurčiček, 2016)

- Goal: NLG based on the syntax tree
  - Encode trees as sequences
  - Seq2Seq model for generation

X is an Italian restaurant near the river.
Contextual NLG (Dušek and Jurčíček, 2016)

- Goal: adapting users’ way of speaking, providing context-aware responses
  - Context encoder
  - Seq2Seq model

![Diagram of Contextual NLG model with prepending context, DA encoder, and decoder with attention.]

Prepending context

DA encoder

Decoder with attention

Context encoder

Contextual additions

Typical NLG

Preceding user utterance

is there another option

Inform(line=M102, direction=Herald Square, vehicle=bus, departure_time=9:01am, from_stop=Wall Street)

Take bus line M102 from Wall Street to Herald Square at 9:01am.

There is a bus at 9:01am from Wall Street to Herald Square using line M102.

textually bound response

Material: [http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
Controlled Text Generation (Hu et al., 2017)

- Idea: NLG based on generative adversarial network (GAN) framework
  - $c$: targeted sentence attributes
NLG Evaluation

- **Metrics**
  - **Subjective: human judgement** (Stent et al., 2005)
    - Adequacy: correct meaning
    - Fluency: linguistic fluency
    - Readability: fluency in the dialogue context
    - Variation: multiple realizations for the same concept
  - **Objective: automatic metrics**
    - Word overlap: BLEU (Papineni et al, 2002), METEOR, ROUGE
    - Word embedding based: vector extrema, greedy matching, embedding average

There is a gap between human perception and automatic metrics
Outline

- PART I. Introduction & Background Knowledge
- **PART II. Task-Oriented Dialogue Systems**
  - Spoken/Natural Language Understanding (SLU/NLU)
  - Dialogue Management – Dialogue State Tracking (DST)
  - Dialogue Management – Dialogue Policy Optimization
  - Natural Language Generation (NLG)
  - *End-to-End Task-Oriented Dialogue Systems*
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
E2E Joint NLU and DM (Yang et al., 2017)

Errors from DM can be propagated to NLU for *regularization + robustness*

Both DM and NLU performance (frame accuracy) is improved
E2E Supervised Dialogue System (Wen et al., 2017)
E2E MemNN for Dialogues (Bordes et al., 2017)

- Split dialogue system actions into subtasks
  - API issuing
  - API updating
  - Option displaying
  - Information informing

<table>
<thead>
<tr>
<th>Task</th>
<th>Memory Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>no match type</td>
</tr>
<tr>
<td>T1: Issuing API calls</td>
<td>99.9 (99.6)</td>
</tr>
<tr>
<td>T2: Updating API calls</td>
<td>100 (100)</td>
</tr>
<tr>
<td>T3: Displaying options</td>
<td>74.9 (2.0)</td>
</tr>
<tr>
<td>T4: Providing information</td>
<td>59.5 (3.0)</td>
</tr>
<tr>
<td>T5: Full dialogs</td>
<td>96.1 (49.4)</td>
</tr>
<tr>
<td>T(00V): Issuing API calls</td>
<td>72.3 (0)</td>
</tr>
<tr>
<td>T2(00V): Updating API calls</td>
<td>78.9 (0)</td>
</tr>
<tr>
<td>T3(00V): Displaying options</td>
<td>74.4 (0)</td>
</tr>
<tr>
<td>T4(00V): Providing inform.</td>
<td>57.6 (0)</td>
</tr>
<tr>
<td>T5(00V): Full dialogs</td>
<td>65.5 (0)</td>
</tr>
<tr>
<td>T6: Dialog state tracking 2</td>
<td>41.1 (0)</td>
</tr>
</tbody>
</table>

Material: [http://opendialogue.miulab.tw](http://opendialogue.miulab.tw)
E2E RL-Based KB-InfoBot (Dhingra et al., 2017)

Entity-Centric Knowledge Base

<table>
<thead>
<tr>
<th>Movie</th>
<th>Actor</th>
<th>Release Year</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Groundhog Day</em></td>
<td>Bill Murray</td>
<td>1993</td>
</tr>
<tr>
<td><em>Australia</em></td>
<td>Nicole Kidman</td>
<td>X</td>
</tr>
<tr>
<td><em>Mad Max: Fury Road</em></td>
<td>X</td>
<td>2015</td>
</tr>
</tbody>
</table>

Idea: differentiable database for propagating the gradients
E2E RL-Based System (Zhao and Eskenazi, 2016)

- Joint learning
  - NLU, DST, Dialogue Policy
- Deep RL for training
  - Deep Q-network
  - Deep recurrent network

Hybrid-RL
Baseline
E2E LSTM-Based Dialogue Control (Williams and Zweig, 2016)

- Idea: an LSTM maps from raw dialogue history directly to a distribution over system actions
  - Developers can provide software including business rules & programmatic APIs → LSTM can take actions in the real world on behalf of the user
  - The LSTM can be optimized using SL or RL
E2E Task-Completion Bot (TC-Bot) (Li et al., 2017)

Text Input
Are there any action movies to see this weekend?

Idea: supervised learning for each component and reinforcement learning for end-to-end training the neural dialogue system
E2E Task-Completion Bot (TC-Bot) \cite{Li2017}

User goal: **Two tickets for “the witch” tomorrow 9:30 PM at regal meridian 16, Seattle.**

**RULE BASED SYSTEM**

User: Can I get 2 tickets for the witch?
Agent: Which theater do you want?
User: Regal meridian 16
Agent: What time would you like to see it?
User: 9:30 pm
Agent: What date would you like to see it?
User: Tomorrow!
Agent: How many people are you?
User: Just two adults.
Agent: What is your location?
User: Seattle!
Agent: What kind of video format do you like?
User: Oh, any kind is okay.
Agent: Do you have any price constraint?
User: No, I don’t care.
Agent: Great - I was able to purchase 2 tickets for you to see the witch tomorrow at regal meridian 16 theater in Seattle at 9:30 pm.
User: Thank you.
Agent: Thank you!

**REINFORCEMENT LEARNING SYSTEM**

User: Can I get 2 tickets for the witch?
Agent: What date would you like to see it?
User: Tomorrow
Agent: Which theater do you prefer?
User: Regal meridian 16
Agent: What is your location?
User: Seattle
Agent: What time would you like to see it?
User: 9:30 pm.
Agent: Great - I was able to purchase 2 tickets for you to see the witch tomorrow at regal meridian 16 theater in Seattle at 9:30 pm.
User: Thanks.
Agent: Thanks!
Hierarchical RL for Composite Tasks (Peng et al., 2017)

- Set of tasks that need to be fulfilled collectively!
- Build a dialog manager that satisfies cross-subtask constraints (slot constraints)
- Temporally constructed goals

- hotel_check_in_time > departure_flight_time
- \# flight_tickets = \#people checking in the hotel
- hotel_check_out_time < return_flight_time,
Hierarchical RL for Composite Tasks (Peng et al., 2017)

- The dialog model makes decisions over two levels: meta-controller and controller.
- The agent learns these policies simultaneously:
  - the policy of optimal sequence of goals to follow \( \pi_g(g_t, s_t; \theta_1) \)
  - Policy \( \pi_{a,g}(a_t, g_t, s_t; \theta_2) \) for each sub-goal \( g_t \)

![Diagram](http://opendialogue.miulab.tw)
Social Chat Bots
Social Chat Bots

- The success of Xiaolce (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+ 16)
  - Persona based models and speaker-role based models (Li+ 16; Luan+ 17)
  - Image-grounded models (Mostafazadeh+ 17)
  - Knowledge-grounded models (Ghazvininejad+ 17)
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
  - Neural Response Generation
  - Response Diversity
  - Response Consistency
  - Deep Reinforcement Learning for Response Generation
  - Combining Task-Oriented Bots and Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Neural Response Generation (Sordoni+ 15; Vinyals & Le 15; Shang+ 15)

Source: conversation history

... because of your game?

encoder

decoder

Target: response

Yeah  I’m  on  my  way
ChitChat Hierarchical Seq2Seq (Serban et al., 2016)

- Learns to generate dialogues from offline dialogs
- No state, action, intent, slot, etc.

Diagram showing the hierarchical Seq2Seq model with encoder, decoder, initial hidden state, context hidden state, and prediction outputs.
ChitChat Hierarchical Seq2Seq (Serban et al., 2017)

- A hierarchical seq2seq model with Gaussian latent variable for generating dialogues (like topic or sentiment)
Neural Response Generation: Blandness Problem

How was your weekend?
I don’t know.

What did you do?
I don’t understand what you are talking about.

This is getting boring...
Yes that’s what I’m saying.

The generated responses are general and meaningless
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
  - Neural Response Generation
  - *Response Diversity*
  - Response Consistency
  - Deep Reinforcement Learning for Response Generation
  - Combining Task-Oriented Bots and Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Mutual Information for Neural Generation (Li et al., 2016)

- Mutual information objective

\[
\hat{T} = \arg \max_T \left\{ \log \frac{p(S, T)}{p(S)p(T)} \right\}
\]

\[
\hat{T} = \arg \max_T \left\{ \log p(T|S) - \lambda \log p(T) \right\}
\]

\[
\hat{T} = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}
\]
Mutual Information for Neural Generation (Li et al., 2016)

\[
\hat{T} = \arg \max_T \left\{ \log p(T|S) - \lambda \log p(T) \right\}
\]

\[
\hat{T}' = \arg \max_T \left\{ (1 - \lambda) \log p(T|S) + \lambda \log p(S|T) \right\}
\]

antiLM: large $\lambda$ makes output ungrammatical

used in rescoring (intractable otherwise)
### Lexical Diversity

# Distinct Tokens in generated targets (divided by total #)

<table>
<thead>
<tr>
<th>System</th>
<th>Distinct Tokens</th>
<th>0.00</th>
<th>0.02</th>
<th>0.04</th>
<th>0.06</th>
<th>0.08</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMT system</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard LSTM</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTM + MMI (antiLM)</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSTM + MMI (bidirectional)</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sample Outputs (Baseline: Maximum Likelihood)

Wow sour starbursts really do make your mouth water... mm drool. Can I have one?

Of course!

Milan apparently selling Zlatan to balance the books... Where next, Madrid?

I don’t know.

‘tis a fine brew on a day like this! Strong though, how many is sensible?

I'm not sure yet, I'll let you know!

Well he was on in Bromley a while ago... still touring.

I don't even know what he's talking about.

32% of the responses:

“I don’t know”

“I don’t know what you are talking about”

“I don’t think that is a good idea”

“Oh my god”

Material: http://opendialogue.miulab.tw
Wow sour starbursts really do make your mouth water... mm drool. Can I have one?

Of course you can! They’re delicious!

Milan apparently selling Zlatan to balance the books... Where next, Madrid?

I think he'd be a good signing.

’tis a fine brew on a day like this! Strong though, how many is sensible?

Depends on how much you drink!

Well he was on in Bromley a while ago... still touring.

I’ve never seen him live.
Sample Outputs (MMI)

Wow sour starbursts really do make your mouth water.. mm drool. Can I have one?

Of course you can! They’re delicious!

Milan apparently selling Zlatan to balance the books... Where next, Madrid?

I think he’d be a good signing.

‘tis a fine brew on a day like this! Strong though, how many is sensible?

Depends on how much you drink!

Well he was on in Bromley a while ago... still touring

I’ve never seen him live.
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
  - Neural Response Generation
  - Response Diversity
  - Response Consistency
  - Deep Reinforcement Learning for Response Generation
  - Combining Task-Oriented Bots and Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Real-World Conversations

- Multimodality
  - Conversation history
  - Persona
  - User profile data (bio, social graph, etc.)
  - Visual signal (camera, picture etc.)
  - Knowledge base
  - Mood
  - Geolocation
  - Time

Because of your game? EOS
Yeah I’m...
Personalized Response Generation (Li et al., 2016)

where do you live in england
Persona Model for Speaker Consistency (Li et al., 2016)

Baseline model:

<table>
<thead>
<tr>
<th>message</th>
<th>Where is your hometown?</th>
</tr>
</thead>
<tbody>
<tr>
<td>response</td>
<td>I live in Indonesia.</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>In which city do you live now?</td>
</tr>
<tr>
<td>response</td>
<td>I live in Jakarta.</td>
</tr>
<tr>
<td>message</td>
<td>In which country do you live now?</td>
</tr>
<tr>
<td>response</td>
<td>I live in Indonesia.</td>
</tr>
<tr>
<td>message</td>
<td>What is your major?</td>
</tr>
<tr>
<td>response</td>
<td>Business, you?</td>
</tr>
<tr>
<td>message</td>
<td>What did you study in college?</td>
</tr>
<tr>
<td>response</td>
<td>Psychology, you?</td>
</tr>
<tr>
<td>message</td>
<td>How old are you?</td>
</tr>
<tr>
<td>response</td>
<td>I’m 18.</td>
</tr>
<tr>
<td>message</td>
<td>What is your age?</td>
</tr>
<tr>
<td>response</td>
<td>I’m 16.</td>
</tr>
</tbody>
</table>

Persona model using speaker embedding [Li+ 16b]

<table>
<thead>
<tr>
<th>message</th>
<th>Where is your hometown?</th>
</tr>
</thead>
<tbody>
<tr>
<td>response</td>
<td>I’m from England.</td>
</tr>
<tr>
<td>message</td>
<td>Where are you from?</td>
</tr>
<tr>
<td>response</td>
<td>I’m from England.</td>
</tr>
<tr>
<td>message</td>
<td>In which city do you live now?</td>
</tr>
<tr>
<td>response</td>
<td>I live in London.</td>
</tr>
<tr>
<td>message</td>
<td>In which country do you live now?</td>
</tr>
<tr>
<td>message</td>
<td>What is your major?</td>
</tr>
<tr>
<td>response</td>
<td>Business, you?</td>
</tr>
<tr>
<td>message</td>
<td>What did you study in college?</td>
</tr>
<tr>
<td>response</td>
<td>I did business studies.</td>
</tr>
<tr>
<td>message</td>
<td>How old are you?</td>
</tr>
<tr>
<td>response</td>
<td>I’m 18.</td>
</tr>
<tr>
<td>message</td>
<td>What is your age?</td>
</tr>
<tr>
<td>response</td>
<td>I’m 18.</td>
</tr>
</tbody>
</table>
Speak-Role Aware Response (Luan et al., 2017)

Speaker independent
Conversational model

Context

Written text

Speaker dependent
Auto encoder model

Response

parameter sharing

Who are you
I ‘m Mary

My name is Mike
My name is Mike
Speak-Role Aware Response (Luan et al., 2017)

- Speaker role: support person
- Domain expertise: technical
- Speaking style: polite

**Context:** have you tried uninstalling nike+ connec? how about docking it to a different computer?
**Input:** Tried both of those techniques. But nothing works. I can try the other computer again.
**Baseline:** Let me know if there’s anything else I can help you with. Thanks for the help.
**Multitask:** Have you tried using a different browser? Let us know if you have any other questions.

**Context:** What do you see when you have a failed login attempt? Never ending spinner? Loop you back to login page?
**Input:** I am getting a loop back to login page, windows.
**Baseline:** Ah, ok. Thanks for the info. I’ll check it out. Thanks for the feedback.
**Multitask:** I’m sorry to hear that. Have you tried clearing your cache and cookies?

**Context:** To clarify: your first run after resetting your connection will not post to Facebook. Does this help? Let us know.
**Input:** That’s correct. I reset it twice! It still doesn’t work.
**Baseline:** Let me know if there’s anything else I can help you with. Thanks for the update.
**Multitask:** I’m sorry to hear that. Are you receiving any error messages? Let us know.
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
  - Neural Response Generation
  - Response Diversity
  - Response Consistency
  - Deep Reinforcement Learning for Response Generation
  - Combining Task-Oriented Bots and Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
Deep RL for Response Generation *(Li et al., 2016)*

- RL agent generates more interactive responses
- RL agent tends to end a sentence with a question and hand the conversation over to the user
- Next step: explore intrinsic rewards, large-scale training

<table>
<thead>
<tr>
<th>Input message</th>
<th>Supervised Learning Agent</th>
<th>Reinforcement Learning Agent</th>
</tr>
</thead>
<tbody>
<tr>
<td>How old are you ?</td>
<td>I’m 16.</td>
<td>I’m 16, why are you asking ?</td>
</tr>
<tr>
<td>What’s your full name ?</td>
<td>I have no idea.</td>
<td>What’s yours ?</td>
</tr>
<tr>
<td>I don’t want to go home tonight.</td>
<td>Really ?</td>
<td>Why ?</td>
</tr>
<tr>
<td>Do you have any feelings for me ?</td>
<td>I don’t know what you are talking about.</td>
<td>I don’t want to hurt your feelings.</td>
</tr>
<tr>
<td>How much time do you have here?</td>
<td>Not long enough. Sorry, sir.</td>
<td>Ten seconds.</td>
</tr>
<tr>
<td>Shall we get started ?</td>
<td>Of course !</td>
<td>Yes. We’ve got a lot of work to do here.</td>
</tr>
<tr>
<td>Do you play football ?</td>
<td>No, i don’t</td>
<td>Yes. I love football !</td>
</tr>
<tr>
<td>We’d have to talk to him.</td>
<td>I mean, he’s a good guy</td>
<td>About what ?</td>
</tr>
<tr>
<td>How come you never say it ?</td>
<td>Because I don’t want to hurt you.</td>
<td>I don’t think it’s a good idea to say it.</td>
</tr>
</tbody>
</table>
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
  - Neural Response Generation
  - Response Diversity
  - Response Consistency
  - Deep Reinforcement Learning for Response Generation
  - *Combining Task-Oriented Bots and Social Chat Bots*
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
The weather is so depressing these days.

I know, I dislike rain too. What about a day trip to eastern Washington?

Any recommendation?

Try Dry Falls, it’s spectacular!
Image-Grounded Dialogues (Mostafazadeh et al., 2017)

Neural model, trained on 250K Twitter convos:
- Turn 1: tweet + accompanying photo
- Turn 2: tweet is a question

Given data, other signals can be exploited in the same way: program state, database info, geolocation, etc.

Tweet: some of the damage done
NeuroCon: Is that at your house?

Tweet: my poor baby moms saying it 's probably totaled so rip Bessie
NeuroCon: How did you manage to do that?

Tweet: as penitence I offer this
NeuroCon: can I have one of those?
Knowledge-Grounded Responses (Ghazvininejad et al., 2017)

Going to Kusakabe tonight

Conversation History

Dialogue Encoder

Σ

Decoder

Fact Encoder

Try omakase, the best in town

Response

Consistently the best omakase
Amazing sushi tasting […]
They were out of kaisui […]

Contextually-Relevant “Facts”

World “Facts”
You know any good **Japanese** restaurant in **Seattle**?

Try **Kisaku**, one of the best **sushi restaurants** in the city.
Knowledge-Grounded Responses (Ghazvininejad et al., 2017)

A: Visiting the celebs at Los Angeles International Airport (LAX) - [...] w/ 70 others
B: Nice airport terminal. Have a safe flight.

A: Is that [...] in your photos? It’s on my list of places to visit in NYC.
B: Don’t forget to check out the 5th floor, while you are here, it’s a great view.

A: Live right now on [...] Tune in!!!!!
B: Listen to Lisa Paige

A: Been craving Chicken Pot Pie who has the best? Trying [...] at [...] Must be Change of weather!
B: Love the pasta trattoria.

A: So [...] is down to one copy of Pound Foolish. I’m curious to see if they are re-ordering it.
B: Check out the video feed on 6 and take a picture of the Simpsons on the 3rd floor.

A: I wish [...] would introduce another vegetarian option besides the shroomburger. It’s delicious but kind of ridiculous.
B: This is the best j.crew in the world. Try the lemonade!

A: Just had an awesome dinner at [...] Great recommendation [...] 
B: One of my favorite places I’ve ever been to in NYC. The food is great and the service is lackluster.

Results (23M conversations) outperforms competitive neural baseline (human + automatic eval)
Evaluation
Dialogue System Evaluation

- Dialogue model evaluation
  - Crowd sourcing
  - User simulator

- Response generator evaluation
  - Word overlap metrics
  - Embedding based metrics
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
  - Human Evaluation
  - User Simulation
  - Objective Evaluation
- PART V. Recent Trends and Challenges
Crowdsourcing for Dialogue System Evaluation (Yang et al., 2012)

The normalized mean scores of Q2 and Q5 for approved ratings in each category. A higher score maps to a higher level of task success.
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
  - Human Evaluation
  - User Simulation
  - Objective Evaluation
- PART V. Recent Trends and Challenges
Goal: generate natural and reasonable conversations to enable reinforcement learning for exploring the policy space

Approach
- Rule-based crafted by experts (Li et al., 2016)
- Learning-based (Schatzmann et al., 2006; El Asri et al., 2016, Crook and Marin, 2017)
Elements of User Simulation

The error model enables the system to maintain the robustness.
Rule-Based Simulator for RL Based System (Li et al., 2016)

- rule-based simulator + collected data
- starts with sets of goals, actions, KB, slot types
- publicly available simulation framework
- movie-booking domain: ticket booking and movie seeking
- provide procedures to add and test own agent
Model-Based User Simulators

- Bi-gram models (Levin et.al. 2000)
- Graph-based models (Scheffler and Young, 2000)
- Data Driven Simulator (Jung et.al., 2009)
- Neural Models (deep encoder-decoder)
Seq2Seq User Simulation (El Asri et al., 2016)

- Seq2Seq trained from dialogue data
  - Input: $c_i$ encodes contextual features, such as the previous system action, consistency between user goal and machine provided values
  - Output: a dialogue act sequence from the user
- Extrinsic evaluation for policy
Seq2Seq User Simulation (Crook and Marin, 2017)

- Seq2Seq trained from dialogue data
  - No labeled data
  - Trained on just human to machine conversations
User Simulator for Dialogue Evaluation Measures

Understanding Ability
- whether constrained values specified by users can be understood by the system
- agreement percentage of system/user understandings over the entire dialog (averaging all turns)

Efficiency
- Number of dialogue turns
- Ratio between the dialogue turns (larger is better)

Action Appropriateness
- an explicit confirmation for an uncertain user utterance is an appropriate system action
- providing information based on misunderstood user requirements
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- **PART IV. Evaluation**
  - Human Evaluation
  - User Simulation
  - *Objective Evaluation*
- PART V. Recent Trends and Challenges
How NOT to Evaluate Dialog System (Liu et al., 2017)

- How to evaluate the quality of the generated response?
  - Specifically investigated for chat-bots
  - Crucial for task-oriented tasks as well

- Metrics:
  - Word overlap metrics, e.g., BLEU, METEOR, ROUGE, etc.
  - Embeddings based metrics, e.g., contextual/meaning representation between target and candidate
Dialogue Response Evaluation (Lowe et al., 2017)

- Problems of existing automatic evaluation
  - can be biased
  - correlate poorly with human judgements of response quality
  - using word overlap may be misleading

- Solution
  - collect a dataset of accurate human scores for variety of dialogue responses (e.g., coherent/un-coherent, relevant/irrelevant, etc.)
  - use this dataset to train an automatic dialogue evaluation model – learn to compare the reference to candidate responses!
  - Use RNN to predict scores by comparing against human scores!

<table>
<thead>
<tr>
<th>Context of Conversation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Speaker A:</strong> Hey, what do you want to do tonight?</td>
</tr>
<tr>
<td><strong>Speaker B:</strong> Why don’t we go see a movie?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Model Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nah, let’s do something active.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reference Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yeah, the film about Turing looks great!</td>
</tr>
</tbody>
</table>
Recent Trends and Challenges

Multimodality
Dialogue Breath
Dialogue Depth
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
  - Multimodality
  - Dialogue Breath
  - Dialogue Depth
Brain Signal for Understanding *(Sridharan et al., 2012)*

- Misunderstanding detection by brain signal
  - Green: listen to the correct answer
  - Red: listen to the wrong answer

Detecting misunderstanding via brain signal in order to correct the understanding results
Video for Intent Understanding

Proactive (from camera)
I want to see a movie on TV!

Intent: turn_on_tv
May I turn on the TV for you?

Proactively understanding user intent to initiate the dialogues.
App Behavior for Understanding (Chen et al., 2015)

- **Task:** user intent prediction
- **Challenge:** language ambiguity

**User preference**
- Some people prefer “Message” to “Email”
- Some people prefer “Ping” to “Text”

**App-level contexts**
- “Message” is more likely to follow “Camera”
- “Email” is more likely to follow “Excel”

Considering behavioral patterns in history to model understanding for intent prediction.
Video Highlight Prediction \((\text{Fu et al., 2017})\)
Video Highlight Prediction \((\text{Fu et al., 2017})\)

- **Goal**: predict highlight from the video
- **Input**: multi-modal and multi-lingual (real time text commentary from fans)
- **Output**: tag if a frame part of a highlight or not
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- **PART V. Recent Trends and Challenges**
  - Multimodality
  - *Dialogue Breath*
  - Dialogue Depth
Evolution Roadmap

Dialogue breadth (coverage)

Dialogue depth (complexity)

I feel sad...

I’ve got a cold what do I do?

Tell me a joke.

What is influenza?
Evolution Roadmap

Dialogue breadth (coverage)

Dialogue depth (complexity)

Single domain systems

Extended systems

Multi-domain systems

Open domain systems

I've got a cold what do I do?

Tell me a joke.

What is influenza?

I feel sad...
Intent Expansion (Chen et al., 2016)

- Transfer dialogue acts across domains
  - Dialogue acts are similar for multiple domains
  - Learning new intents by information from other domains

The dialogue act representations can be automatically learned for other domains.
Policy for Domain Adaptation (Gašić et al., 2015)

- Bayesian committee machine (BCM) enables estimated Q-function to share knowledge across domains

The policy from a new domain can be boosted by the committee policy
Outline

- PART I. Introduction & Background Knowledge
- PART II. Task-Oriented Dialogue Systems
- PART III. Social Chat Bots
- PART IV. Evaluation
- PART V. Recent Trends and Challenges
  - Multimodality
  - Dialogue Breath
  - Dialogue Depth
Evolution Roadmap

Dialogue breadth (coverage)

Dialogue depth (complexity)

Empathetic systems

I feel sad...

Common sense system

Tell me a joke.

What is influenza?

Knowledge based system

I’ve got a cold what do I do?

What is influenza?
High-Level Intention for Dialogue Planning (Sun et al., 2016)

- High-level intention may span several domains

Users can interact via high-level descriptions and the system learns how to plan the dialogues.
Empathy in Dialogue System (Fung et al., 2016)

- Embed an empathy module
  - Recognize emotion using multimodality
  - Generate emotion-aware responses

Emotion Recognizer

Text

Speech

Vision

Emotion Recognizer

Material: http://opendialogue.miulab.tw
Visual Object Discovery through Dialogues (Vries et al., 2017)

- Recognize objects using “Guess What?” game
- Includes “spatial”, “visual”, “object taxonomy” and “interaction”
Conclusion
Summarized Challenges

- Human-machine interfaces is a hot topic but building a good one is challenging!
- Most state-of-the-art technologies are based on DNN
  - Requires huge amounts of labeled data
  - Several frameworks/models are available
- Leveraging structured knowledge and unstructured data
- Handling reasoning
- Data collection and analysis from un-structured data
- The capability of task-oriented and chit-chat dialogues should be integrated.
Brief Conclusions

- Introduce recent deep learning methods used in dialogue models
- Highlight main components of task-oriented dialogue systems and new deep learning architectures used for these components
- Highlight the challenges and trends for current chat bot research
- Talk about new avenues for current state-of-the-art dialogue research
- Provide all materials online!

http://opendialogue.miulab.tw
Thanks to Dilek Hakkani-Tur, Asli Celikyilmaz, Tsung-Hsien Wen, Pei-Hao Su, Li Deng, Sungjin Lee, Milica Gašić, Lihong Li, Xiujin Li, Abhinav Rastogi, Ankur Bapna, Pararth Shah and Gokhan Tur for sharing their slides.

THANKS FOR ATTENTION!