人工智慧任務導向對話系統 (AI Task-Oriented Dialogue System)

中國文化大學應用數學系專題演講
Host: Prof. Chia-Hui Liu
Department of Applied Mathematics, Chinese Culture University
Time: 13:10-15:00, Nov 27, 2020 (Friday)
Place: 理學院會議室 (大義館610), CCU
Address: 55, Hwa-Kang Rd., Yang-Ming-Shan, Taipei, Taiwan

Min-Yuh Day
戴敏育
Associate Professor
副教授
Institute of Information Management, National Taipei University
國立台北大學 資訊管理研究所

https://web.ntpu.edu.tw/~myday
2020-11-27
戴敏育博士
(Min-Yuh Day, Ph.D.)

國立台北大學 資訊管理研究所 副教授
中央研究院 資訊科學研究所 訪問學人
國立台灣大學 資訊管理 博士

Publications Co-Chairs, IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013-)

Program Co-Chair, IEEE International Workshop on Empirical Methods for Recognizing Inference in TExt (IEEE EM-RITE 2012-)

Publications Chair, The IEEE International Conference on Information Reuse and Integration (IEEE IRI)
Outline

• Artificial Intelligence
• Question Answering
• Dialogue Systems
• Task Oriented Dialogue System
AIWISFIN
AI Conversational Robo-Advisor
(人工智能對話式理財機器人)
First Place, InnoServe Awards 2018

https://www.youtube.com/watch?v=sEhmyoTXmGk
2018 The 23<sup>th</sup> International ICT Innovative Services Awards (InnoServe Awards 2018)

- Annual ICT application competition held for university and college students
- The largest and the most significant contest in Taiwan.
- More than ten thousand teachers and students from over one hundred universities and colleges have participated in the Contest.

https://innoserve.tca.org.tw/award.aspx
2018 International ICT Innovative Services Awards (InnoServe Awards 2018)
(2018 第23屆大專校院資訊應用服務創新競賽)

https://innoserve.tca.org.tw/award.aspx
AI
Artificial Intelligence (A.I.) Timeline

1950
TURING TEST
Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence.

1955
A.I. BORN
Term ‘artificial intelligence’ is coined by computer scientist, John McCarthy to describe “the science and engineering of making intelligent machines”

1961
UNIMATE
First industrial robot, Unimate, goes to work at GM replacing humans on the assembly line

1964
ELIZA
Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans

1966
SHAKEY
The ‘first electronic person’ from Stanford, Shakey is a general-purpose mobile robot that reasons about its own actions

A.I. WINTER
Many false starts and dead-ends leave A.I. out in the cold

1997
DEEP BLUE
Deep Blue, a chess-playing computer from IBM defeats world chess champion Garry Kasparov

1998
KISMET
Cynthia Breazeal at MIT introduces Kismet, an emotionally intelligent robot insofar as it detects and responds to people’s feelings

1999
AIBO
Sony launches first consumer robot pet dog AIBO (A.I. robot) with skills and personality that develop over time

2002
ROOMBA
First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes

2011
SIRI
Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S

2011
WATSON
IBM’s question answering computer Watson wins first place on popular $1M prize television quiz show Jeopardy

2014
EUGENE
Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human

2014
ALEXA
Amazon launches Alexa, an intelligent virtual assistant with a voice interface that completes shopping tasks

2016
TAY
Microsoft’s chatbot Tay goes rogue on social media making inflammatory and offensive racist comments

2017
ALPHAGO
Google’s A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2170) of possible positions

The Rise of AI

1.1 Origin & Definition of AI

Artificial intelligence (AI) is not new. The term was coined in 1956 by John McCarthy, a Stanford computer science professor who organized an academic conference on the topic at Dartmouth College in the summer of that year. The field of AI has gone through a series of boom-bust cycles since then, characterized by technological breakthroughs that stirred activity and excitement about the topic, followed by subsequent periods of disillusionment and disinterest known as ‘AI Winters’ as technical limitations were discovered. As you can see in figure 1, today we are once again in an ‘AI Spring’.

Artificial intelligence can be defined as human intelligence exhibited by machines; systems that approximate, mimic, replicate, automate, and eventually improve on human thinking. Throughout the past half-century a few key components of AI were established as essential: the ability to perceive, understand, learn, problem solve, and reason. Countless working definitions of AI have been proposed over the years but the unifying thread in all of them is that computers with the right software can be used to solve the kind of problems that humans solve, interact with humans and the world as humans do, and create ideas like humans. In other words, while the mechanisms that give rise to AI are ‘artificial’, the intelligence to which AI is intended to approximate is indistinguishable from human intelligence. In the early days of the science, processing inputs from the outside world required extensive programming, which limited early AI systems to a very narrow set of inputs and conditions. However since then, computer science has worked to advance the capability of AI-enabled computing systems.

Board games have long been a proving ground for AI research, as they typically involve a finite number of players, rules, objectives, and possible moves. This essentially means that games – one by one, including checkers, backgammon, and even Jeopardy! to name a few – have been taken over by AI. Most famously, in 1997 IBM’s Deep Blue defeated Garry Kasparov, the then reigning world champion of chess. This trajectory persists with the ancient Chinese game of Go, and the defeat of reigning world champion Lee Sedol by DeepMind’s AlphaGo in March 2016.

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1956</td>
<td>Dartmouth conference led by John McCarthy coins the term “artificial intelligence”</td>
</tr>
<tr>
<td>1964</td>
<td>Eliza, the first chatbot is developed by Joseph Weizenbaum at MIT</td>
</tr>
<tr>
<td>1975 – 1982</td>
<td>Edward Feigenbaum develops the first Expert System, giving rebirth to AI</td>
</tr>
<tr>
<td>1997</td>
<td>IBM’s Deep Blue defeats Garry Kasparov, the world’s reigning chess champion</td>
</tr>
<tr>
<td>2011</td>
<td>IBM’s Watson Q&amp;A machine wins Jeopardy!</td>
</tr>
<tr>
<td>2014</td>
<td>Apple integrates Siri, a personal voice assistant into the iPhone</td>
</tr>
<tr>
<td>2016</td>
<td>AlphaGo defeats Lee Sedol</td>
</tr>
<tr>
<td>2020</td>
<td>YouTube recognizes cats from videos</td>
</tr>
</tbody>
</table>

**AI Winter I**
- Limited computer processing power
- Limited database storage capacity
- Limited network ability
- Real-world problems are complicated
- Facial recognition, translation
- Combinatorial explosion
- Disappointing results: failure to achieve scale
- Collapse of dedicated hardware vendors

**AI Winter II**
- Machine learning
- Deep learning: pattern analysis & classification
- Big data: large databases
- Fast processors to crunch data
- High-speed networks and connectivity

Definition of Artificial Intelligence (A.I.)
Artificial Intelligence

“... the science and engineering of making intelligent machines”

(John McCarthy, 1955)
Artificial Intelligence

“... technology that thinks and acts like humans”

Artificial Intelligence

“... intelligence exhibited by machines or software”

## 4 Approaches of AI

<table>
<thead>
<tr>
<th>Thinking Humanly</th>
<th>Thinking Rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acting Humanly</td>
<td>Acting Rationally</td>
</tr>
</tbody>
</table>

|--------------------------------------------------|-----------------------------------------------|

AI Acting Humanly: The Turing Test Approach (Alan Turing, 1950)

- Knowledge Representation
- Automated Reasoning
- Machine Learning (ML)
  - Deep Learning (DL)
- Computer Vision (Image, Video)
- Natural Language Processing (NLP)
- Robotics

IMTKU
Emotional Dialogue System for Short Text Conversation at NTCIR-14 STC-3 (CECG) Task
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-9 RITE

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Chun Tu
myday@mail.tku.edu.tw

NTCIR-9 Workshop, December 6-9, 2011, Tokyo, Japan
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-10 RITE-2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Chun Tu
Hou-Cheng Vong
Shih-Wei Wu
Shih-Jhen Huang

myday@mail.tku.edu.tw
IMTKU Textual Entailment System for Recognizing Inference in Text at NTCIR-11 RITE-VAL

Tamkang University

2014

Min-Yuh Day
Ya-Jung Wang
Che-Wei Hsu
En-Chun Tu

Huai-Wen Hsu
Yu-An Lin
Shang-Yu Wu
Yu-Hsuan Tai
Cheng-Chia Tsai

NTCIR-11 Conference, December 8-12, 2014, Tokyo, Japan
IMTKU Question Answering System for World History Exams at NTCIR-12 QA Lab2

Department of Information Management
Tamkang University, Taiwan

Min-Yuh Day
Cheng-Chia Tsai
Wei-Chun Chung
Hsiu-Yuan Chang
Tzu-Jui Sun
Yuan-Jie Tsai
Jin-Kun Lin
Cheng-Hung Lee
Yu-Ming Guo
Yue-Da Lin
Wei-Ming Chen
Yun-Da Tsai
Cheng-Jhih Han
Yi-Jing Lin
Yi-Heng Chiang
Ching-Yuan Chien

myday@mail.tku.edu.tw

NTCIR-12 Conference, June 7-10, 2016, Tokyo, Japan
IMTKU Question Answering System for World History Exams at NTCIR-13 QALab-3

Department of Information Management
Tamkang University, Taiwan

2017

myday@mail.tku.edu.tw

NTCIR-13 Conference, December 5-8, 2017, Tokyo, Japan
IMTKU Emotional Dialogue System for Short Text Conversation at NTCIR-14 STC-3 (CECG) Task

Department of Information Management
Tamkang University, Taiwan

myday@mail.tku.edu.tw

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
### 2020 NTCIR-15 Dialogue Evaluation (DialEval-1) Task

**Dialogue Quality (DQ) and Nugget Detection (ND)**

**Chinese Dialogue Quality (S-score) Results** *(Zeng et al., 2020)*

<table>
<thead>
<tr>
<th>Run</th>
<th>Mean RSNOD</th>
<th>Run</th>
<th>Mean NMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMTKU-run2</td>
<td>0.1918</td>
<td>IMTKU-run2</td>
<td>0.1254</td>
</tr>
<tr>
<td>IMTKU-run1</td>
<td>0.1964</td>
<td>IMTKU-run0</td>
<td>0.1284</td>
</tr>
<tr>
<td>IMTKU-run0</td>
<td>0.1977</td>
<td>IMTKU-run1</td>
<td>0.1290</td>
</tr>
<tr>
<td>TUA1-run2</td>
<td>0.2024</td>
<td>TUA1-run2</td>
<td>0.1310</td>
</tr>
<tr>
<td>TUA1-run0</td>
<td>0.2053</td>
<td>TUA1-run0</td>
<td>0.1322</td>
</tr>
<tr>
<td>NKUST-run1</td>
<td>0.2057</td>
<td>NKUST-run1</td>
<td>0.1363</td>
</tr>
<tr>
<td>BL-lstm</td>
<td>0.2088</td>
<td>BL-lstm</td>
<td>0.1397</td>
</tr>
<tr>
<td>WUST-run0</td>
<td>0.2131</td>
<td>BL-popularity</td>
<td>0.1442</td>
</tr>
<tr>
<td>RSLNV-run0</td>
<td>0.2141</td>
<td>RSLNV-run0</td>
<td>0.1483</td>
</tr>
<tr>
<td>BL-popularity</td>
<td>0.2288</td>
<td>TUA1-run1</td>
<td>0.1540</td>
</tr>
<tr>
<td>TUA1-run1</td>
<td>0.2302</td>
<td>WUST-run0</td>
<td>0.1540</td>
</tr>
<tr>
<td>NKUST-run0</td>
<td>0.2653</td>
<td>NKUST-run0</td>
<td>0.2289</td>
</tr>
<tr>
<td>BL-uniform</td>
<td>0.2811</td>
<td>BL-uniform</td>
<td>0.2497</td>
</tr>
</tbody>
</table>

IMTKU System Architecture for NTCIR-13 QALab-3

- Question Analysis
- Document Retrieval
- Answer Extraction
- Answer Generation

Inputs:
- Question (XML)
- Complex Essay
- Simple Essay
- True-or-False
- Factoid
- Slot-Filling
- Unique

External Resources:
- JA&EN Translator
- Stanford CoreNLP
- Wikipedia

Word Embedding:
- Wiki Word2Vec
System Architecture of Intelligent Dialogue and Question Answering System

- User Question Input
- Dialogue Intention Detection
- AIML Dialogue Engine
- Real Time Dialogue API
- System Response Generator
- Question Analysis
- Document Retrieval
- Answer Extraction
- Answer Generation
- Answer Validation
- Deep Learning
- TensorFlow
- Python NLTK
- Dialogue KB
- IR
- Deep Learning

- RNN
- LSTM
- GRU
- AIML KB
- Cloud Resource
IMTKU Emotional Dialogue
System Architecture

1. Retrieval-Based Model
2. Generation-Based Model
3. Emotion Classification Model
4. Response Ranking

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
The system architecture of IMTKU retrieval-based model for NTCIR-14 STC-3

Retrieval-Based Model

- Post
  - Word Segmentation
- Corpus
  - Building Index
  - Keyword Boolean Query
  - Solr Matching
  - Distinct Result Data
  - Emotion Matching
  - Emotion Classification
  - Retrieval Model
  - Word2Vec Similarity Ranking
  - Retrieval-Based Response

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
The system architecture of IMTKU generation-based model for NTCIR-14 STC-3

Generation-Based Model

Training Data

Building Word Index

Word Embedding

Training Data Seq2seq model

Post

Word Segmentation

Short Text Emotion Classifier

Trained Model

Emotion Matching

Word2Vec Similarity Ranking

Generation-Based Response

Generative Model
The system architecture of IMTKU emotion classification model for NTCIR-14 STC-3

Emotion Classification Model

Corpus → Emotion Classification

Emotion Classification Model

Training Dataset → MLP

LSTM BiLSTM

Testing Dataset

Emotion Classification Model

Emotion Prediction

NTCIR-14 Conference, June 10-13, 2019, Tokyo, Japan
The system architecture of IMTKU Response Ranking for NTCIR-14 STC-3

Response Ranking

1. STC3 Corpus
2. Chinese Segmentation using Jieba
3. Stop Words Removal
4. Word2Vec
5. 1.2 million data (300 dimensions)
6. Vector of Corpus
Short Text Conversation Task (STC-3)  
Chinese Emotional Conversation Generation (CECG) Subtask
# NTCIR Short Text Conversation

## STC-1, STC-2, STC-3

<table>
<thead>
<tr>
<th></th>
<th>Japanese</th>
<th>Chinese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NTCIR-12 STC-1</strong>&lt;br&gt;22 active participants</td>
<td>Twitter, Retrieval</td>
<td>Weibo, Retrieval</td>
<td></td>
</tr>
<tr>
<td><strong>NTCIR-13 STC-2</strong>&lt;br&gt;27 active participants</td>
<td>Yahoo! News, Retrieval+ Generation</td>
<td>Weibo, Retrieval+ Generation</td>
<td></td>
</tr>
<tr>
<td><strong>NTCIR-14 STC-3</strong></td>
<td></td>
<td></td>
<td>Weibo, Generation for given emotion categories</td>
</tr>
<tr>
<td><strong>Chinese Emotional Conversation Generation (CECG) subtask</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dialogue Quality (DQ) and Nugget Detection (ND) subtasks</strong></td>
<td></td>
<td></td>
<td>Weibo+English translations, distribution estimation for subjective annotations</td>
</tr>
</tbody>
</table>

Source: [https://waseda.app.box.com/v/STC3atNTCIR-14](https://waseda.app.box.com/v/STC3atNTCIR-14)

- **Single-turn, Non task-oriented**
- **Multi-turn, task-oriented (helpdesk)**
Chatbots: Evolution of UI/UX

<table>
<thead>
<tr>
<th>Paradigm</th>
<th>mid - 80s</th>
<th>mid - 90s</th>
<th>mid - 00s</th>
<th>mid - 10s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>PC</td>
<td>Web</td>
<td>Smartphone</td>
<td>Messaging</td>
</tr>
<tr>
<td>Examples</td>
<td>Desktop DOS, Windows, Mac OS</td>
<td>Browser Mosaic, Explorer, Chrome</td>
<td>Mobile OS iOS, Android</td>
<td>Messaging Apps WhatsApp, Messenger, Slack</td>
</tr>
<tr>
<td>Applications</td>
<td>Clients Excel, PPT, Lotus</td>
<td>Website Yahoo, Amazon</td>
<td>Apps Angry Birds, Instagram</td>
<td>Bots Weather, Travel</td>
</tr>
<tr>
<td>UI/UX</td>
<td>Native Screens</td>
<td>Web Pages</td>
<td>Native Mobile Screens</td>
<td>Message</td>
</tr>
<tr>
<td>S/w Dev</td>
<td>Client-side</td>
<td>Server-side</td>
<td>Client-side</td>
<td>Server-side</td>
</tr>
</tbody>
</table>

Source: https://bbvaopen4u.com/en/actualidad/want-know-how-build-conversational-chatbot-here-are-some-tools
AI Humanoid
Robo-Advisor
AI Humanoid Robo-Advisor for Multi-channel Conversational Commerce

AI Portfolio Asset Allocation

AI Conversation Dialog System

Multichannel Platforms
- Web
- LINE
- Facebook
- Humanoid Robot
System Architecture of AI Humanoid Robo-Advisor
Conversational Model
(LINE, FB Messenger)
Conversational Robo-Advisor
Multichannel UI/UX
Robots

ALPHA 2

ZENBO
AI Dialogue System
Dialogue Subtasks

Dialogue subtasks

- **Dialogue Generation**
  - 9 leaderboards
  - 35 papers with code

- **Dialogue State Tracking**
  - 2 leaderboards
  - 30 papers with code

- **Visual Dialog**
  - 3 leaderboards
  - 28 papers with code

- **Task-Oriented Dialogue Systems**
  - 20 papers with code

- **Goal-Oriented Dialog**
  - 15 papers with code

- **Short-Text Conversation**
  - 6 papers with code

Source: [https://paperswithcode.com/area/natural-language-processing/dialogue](https://paperswithcode.com/area/natural-language-processing/dialogue)
Chatbot
Dialogue System
Intelligent Agent
Dialogue System

Overall Architecture of Intelligent Chatbot

Can machines think?

(Alan Turing, 1950)

Chatbot

“online human-computer dialog system with natural language.”

Chatbot Conversation Framework

- **Open Domain**: Impossible
- **Closed Domain**: Rules-Based [Easiest]
- **Generative-Based**: Smart Machine [Hard]
- **Retrieval-Based**: General AI [Hardest]

Source: https://chatbotslife.com/ultimate-guide-to-leveraging-nlp-machine-learning-for-you-chatbot-531ff2dd870c
Chatbots

Bot Maturity Model

Customers want to have simpler means to interact with businesses and get faster response to a question or complaint.

From E-Commerce to Conversational Commerce: Chatbots and Virtual Assistants

Source: http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/
Conversational Commerce: eBay AI Chatbots

Hotel Chatbot

**BookHotel**

- I’d like to book a hotel
- Sure, which city?
- New York City
- What date are you leaving?
- November 30th, 2016
- Are you sure you want to book the hotel in NYC?
- Yes
- Thank you. The reservation went through successfully.

**Intents**
An intent performs an action in response to natural language user input.

**Utterances**
Spoken or typed phrases that invoke your intent.

**Slots**
Slots are input data required to fulfill the intent.

**Fulfillment**
Fulfillment mechanism for your intent.

Source: https://sdtimes.com/amazon/guest-view-capitalize-amazon-lex-available-general-public/
H&M’s Chatbot on Kik
Uber’s Chatbot on Facebook’s Messenger

- one main benefit: it loads much faster than the Uber app

Source: http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/
Bot Life Cycle and Platform Ecosystem
The Bot Lifecycle

Source: https://chatbotsmagazine.com/the-bot-lifecycle-1ff357430db7
The bot platform ecosystem and the emerging giants

Nearly every large software company has announced some sort of bot strategy in the last year. Here's a look at a handful of leading platforms that developers might use to send messages, interpret natural language, and deploy bots, with the emerging bot-ecosystem giants highlighted.

General AI agents with platforms
Developer access available now or announced

Source: https://www.oreilly.com/ideas/infographic-the-bot-platform-ecosystem
Bots Landscape

Connectors/Shared Services

AI Tools: Natural Language Processing, Machine Learning, Speech & Voice Recognition

Bot Discovery

Bot developer frameworks and tools

Analytics

Messaging

Messenger Bot Landscape

May 2017

Food
- The Wise Parent
- Pharm
- Perscription Kitchen
- Hungry
- Foodie
- Fitmeal
- Bannock
- Chatoorkut
- Make My Sushi
- Yummy

Communication
- TangoNews
- Typeform
- Accely
- TuneMy
- Ringgo
- Rescue
- Messenger Match
- Sense
- LangLearningBot
- Chat Club
- Umgio Translate
- Swedbank
- UhReport Global
- Trigger

Utilities
- Pancho
- CardBot
- Snokey
- DotCom
- Server Monitoring
- English Dictionary
- Youtube Search
- View Bot
- EllRobot
- Instant Translator

Personal
- M
- Assistant
- Operator
- Uber
- Sniply
- AskWills
- Neo Build
- SelectionBot
- Bud Light Ice
- AskGaryVee
- Guy
- Vivial

Analytics
- SISENSE
- StockTrek
- Page Insights
- DMM
- BuzzLogger
- Trading Bot

Design
- ColorBot
- Connie Digital
- AWAY
- Mr. Norman
- Graphic Design
- SnapBot

News
- CNN
- TMZ
- Digg
- WSJ
- Reddit Bot
- Al-Jazeera
- Hacker News
- Wired
- The Guardian
- France Info
- Chatbots Mag
- VentureBeat

Travel
- Gitease
- KLM
- British Airways
- Space Explorer
- Austrian Airlines
- SnapTravel
- Skyscanner
- Kayak
- TicketBot
- Rapidio

Entertainment
- Spotify
- Kim Kardashian
- La Bilingual
- 50 Cent
- Caquillo Ful
- Lindsey Lohan
- Marcos S
- MTV News
- Ansel Elgort
- Red Bull TV
- Surround
- Star Wars Bot
- Cifem
- Poketbot

Developer Tools
- HackerOne
- Winedeltra
- Rebbie
- Zify
- MemortexBot
- Kinschi

Education
- Genius
- Kindchi

Source: https://medium.com/@RecastAI/2017-messenger-bot-landscape-a-public-spreadsheet-gathering-1000-messenger-bots-f017fdb1448a
How to Build Chatbots

Chatbot Frameworks and AI Services

• Bot Frameworks
  – Botkit
  – Microsoft Bot Framework
  – Rasa NLU

• AI Services
  – Wit.ai
  – api.ai
  – LUIS.ai
  – IBM Watson

# Chatbot Frameworks

## Comparison Table of Most Prominent Bot Frameworks

<table>
<thead>
<tr>
<th>Feature</th>
<th>Botkit</th>
<th>Microsoft Bot Framework</th>
<th>RASA NLU</th>
</tr>
</thead>
<tbody>
<tr>
<td>Built-in integration with messaging platforms</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td>NLP support</td>
<td>✗ but possible to integrate with middlewares</td>
<td>✗ but have close bonds with LUIS.ai</td>
<td>✔️</td>
</tr>
<tr>
<td>Out-of-box bots ready to be deployed</td>
<td>✔️</td>
<td>❌</td>
<td>❌</td>
</tr>
<tr>
<td>Programming Language</td>
<td>JavaScript (Node)</td>
<td>JavaScript (Node), C#</td>
<td>Python</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>wit.ai</th>
<th>api.ai</th>
<th>LUIS.ai</th>
<th>IBM Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free of charge</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> but has paid enterprise version</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> it is in beta and has transaction limits</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> 30 days trial then priced for enterprise use</td>
</tr>
<tr>
<td>Text and Speech processing</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> with use of Cortana</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
</tr>
<tr>
<td>Machine Learning Modeling</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
</tr>
<tr>
<td>Support for Intents, Entities, Actions</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> Intents used as trait entities, actions are combined operations</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> Intents is the main prediction mechanism. Domains of entities, intents and actions</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
</tr>
<tr>
<td>Pre-build entities for easy parsing of numbers, temperature, date, etc.</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /></td>
</tr>
<tr>
<td>Integration to messaging platforms</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✗" /> web service API</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> also has facility for deploying to heroku. Paid environment</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> integrated to Azure</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> possible via API</td>
</tr>
<tr>
<td>Support of SDKs</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> includes SDKs for Python, Node.js, Rust, C, Ruby, iOS, Android, Windows Phone</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> C#, Xamarin, Python, Node.js, iOS, Android, Windows Phone</td>
<td><img src="https://i.imgur.com/1Q5.jpg" alt="✓" /> enables building with Web Service API, Microsoft Bot Framework integration</td>
<td>Proprietary language “AlchemyLanguage”</td>
</tr>
</tbody>
</table>

Transformer (Attention is All You Need) (Vaswani et al., 2017)

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

Overall pre-training and fine-tuning procedures for BERT

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

BERT input representation

BERT, OpenAI GPT, ELMo

Fine-tuning BERT on Different Tasks

(a) Sentence Pair Classification Tasks: MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG

(b) Single Sentence Classification Tasks: SST-2, CoLA

(c) Question Answering Tasks: SQuAD v1.1

(d) Single Sentence Tagging Tasks: CoNLL-2003 NER

Fine-tuning BERT on Question Answering (QA)

(c) Question Answering Tasks: SQuAD v1.1

Fine-tuning BERT on Dialogue Intent Detection (ID; Classification)

Fine-tuning BERT on Dialogue Slot Filling (SF)

Transformers
State-of-the-art Natural Language Processing for TensorFlow 2.0 and PyTorch

- Transformers
  - pytorch-transformers
  - pytorch-pretrained-bert
- provides state-of-the-art general-purpose architectures
  - (BERT, GPT-2, RoBERTa, XLM, DistilBert, XLNet, CTRL...)
  - for Natural Language Understanding (NLU) and Natural Language Generation (NLG)
  - with over 32+ pretrained models in 100+ languages and deep interoperability between TensorFlow 2.0 and PyTorch.

Source: https://github.com/huggingface/transformers
Transfer Learning in Natural Language Processing

Question Answering (QA)
SQuAD
Stanford Question Answering Dataset
What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or span, from the corresponding reading passage, or the question might be unanswerable.

SQuAD2.0 combines the 100,000 questions in SQuAD 1.1 with over 50,000 unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering.

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Model</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Human Performance</td>
<td>86.831</td>
<td>89.452</td>
</tr>
<tr>
<td></td>
<td>Stanford University</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Rajpurkar &amp; Jia et al. ‘18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>SA-Net on Albert (ensemble)</td>
<td>90.724</td>
<td>93.011</td>
</tr>
<tr>
<td></td>
<td>QIANXIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>SA-Net-V2 (ensemble)</td>
<td>90.679</td>
<td>92.948</td>
</tr>
<tr>
<td></td>
<td>QIANXIN</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Retro-Reader (ensemble)</td>
<td>90.578</td>
<td>92.978</td>
</tr>
</tbody>
</table>

https://rajpurkar.github.io/SQuAD-explorer/
SQuAD: 100,000+ Questions for Machine Comprehension of Text

Pranav Rajpurkar and Jian Zhang and Konstantin Lopyrev and Percy Liang
{pranavsr,zjian,klopyrev,pliang}@cs.stanford.edu
Computer Science Department
Stanford University

Abstract

We present the Stanford Question Answering Dataset (SQuAD), a new reading comprehension dataset consisting of 100,000+ questions posed by crowdworkers on a set of Wikipedia articles, where the answer to each question is a segment of text from the corresponding reading passage. We analyze the dataset to understand the types of reasoning required to answer the questions, leaning heavily on dependency and constituency trees. We build a strong logistic regression model, which achieves an F1 score of 51.0%, a significant improvement over a simple baseline (20%). However, human performance (86.8%) is much higher, indicating that the dataset presents a good challenge problem for future research. The dataset is freely available at https://stanford-qa.com.

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

What causes precipitation to fall?
gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
graupel

Where do water droplets collide with ice crystals to form precipitation?
within a cloud

Figure 1: Question-answer pairs for a sample passage in the
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity from clouds. The main forms of precipitation include drizzle, rain, sleet, snow, ice pellets, graupel and hail. Precipitation occurs when a portion of the atmosphere becomes saturated with water vapor (reaching 100% relative humidity), so that the water condenses and "precipitates". Thus, fog and mist are not precipitation but suspensions, because the water vapor does not condense sufficiently to precipitate. Two processes, possibly acting together, can lead to air becoming saturated: cooling the air or adding water vapor to the air. Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called "showers."
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?
A: gravity
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

**Q:** What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?

**A:** graupel
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called “showers”.

**Q:** Where do water droplets collide with ice crystals to form precipitation?

**A:** **within a cloud**
In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail...

Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?
A: gravity

Q: What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?
A: graupel

Q: Where do water droplets collide with ice crystals to form precipitation?
A: within a cloud
Super Bowl 50 was an American football game to determine the champion of the National Football League (NFL) for the 2015 season. The American Football Conference (AFC) champion Denver Broncos defeated the National Football Conference (NFC) champion Carolina Panthers, 24–10. The game was played on February 7, 2016, at Levi's Stadium in Santa Clara, California, in the San Francisco Bay Area. As this was the 50th Super Bowl game, the league emphasized the "golden anniversary" with various gold-themed initiatives during the 2015 season, as well as suspending the tradition of naming each Super Bowl game with Roman numerals (under which the game would have been known as "Super Bowl L"), so the logo could prominently feature the Arabic numerals 5 and 0.[5][6]

The Panthers finished the regular season with a 15–1 record, ranking up the league’s top offense, and quarterback Cam Newton was named the NFL Most Valuable Player (MVP). They defeated the Arizona Cardinals 49–15 in the NFC Championship Game and advanced to their second Super Bowl appearance since the franchise began playing in 1995. The Broncos finished the regular season with a 12–4 record, bolstered by having the league’s top defense. The Broncos defeated the defending Super Bowl champion New England Patriots 20–18 in the AFC Championship Game joining the Patriots, Dallas Cowboys, and Pittsburgh Steelers as one of four teams that have made eight appearances in the Super Bowl. This record would later be broken the next season, in 2017, when the Patriots advanced to their ninth Super Bowl appearance in Super Bowl LI.

https://en.wikipedia.org/wiki/Super_Bowl_50
Dialogue on Airline Travel Information System (ATIS)
The ATIS (Airline Travel Information System) Dataset

https://www.kaggle.com/siddhadev/atis-dataset-from-ms-cntk

<table>
<thead>
<tr>
<th>Sentence</th>
<th>what</th>
<th>flights</th>
<th>leave</th>
<th>from</th>
<th>phoenix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slots</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>B-fromloc</td>
</tr>
<tr>
<td>Intent</td>
<td></td>
<td></td>
<td>atis_flight</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Training samples: 4978
Testing samples: 893
Vocab size: 943
Slot count: 129
Intent count: 26

SF-ID Network (E et al., 2019)
Slot Filling (SF)
Intent Detection (ID)

A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling

Intent Detection on ATIS
State-of-the-art

### Table: State-of-the-art methods

<table>
<thead>
<tr>
<th>RANK</th>
<th>METHOD</th>
<th>ACCURACY</th>
<th>PAPER TITLE</th>
<th>YEAR</th>
<th>PAPER</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SF-ID</td>
<td>0.9776</td>
<td>A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling</td>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Capsule-NLU</td>
<td>0.950</td>
<td>Joint Slot Filling and Intent Detection via Capsule Neural Networks</td>
<td>2018</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: https://paperswithcode.com/sota/intent-detection-on-atis
Slot Filling on ATIS
State-of-the-art

Slot Filling on ATIS

<table>
<thead>
<tr>
<th>RANK</th>
<th>METHOD</th>
<th>F1</th>
<th>PAPER TITLE</th>
<th>YEAR</th>
<th>PAPER</th>
<th>CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SF-ID</td>
<td>0.958</td>
<td>A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling</td>
<td>2019</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Capsule-NLU</td>
<td>0.952</td>
<td>Joint Slot Filling and Intent Detection via Capsule Neural Networks</td>
<td>2018</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: https://paperswithcode.com/sota/slot-filling-on-atis
Restaurants Dialogue Datasets

• MIT Restaurant Corpus
  – https://groups.csail.mit.edu/sls/downloads/restaurant/

• CamRest676
  (Cambridge restaurant dialogue domain dataset)
  – https://www.repository.cam.ac.uk/handle/1810/260970

• DSTC2 (Dialog State Tracking Challenge 2 & 3)
  – http://camdial.org/~mh521/dstc/
### CrossWOZ: A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

<table>
<thead>
<tr>
<th>MultiWOZ</th>
<th>CrossWOZ</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>usr:</strong> I’m looking for a college type attraction.</td>
<td><strong>usr:</strong> 你好，可以帮我推荐一个评分是4.5分以上的景点吗？</td>
</tr>
<tr>
<td></td>
<td>Hello, could you recommend an attraction with a rating of 4.5 or higher?</td>
</tr>
<tr>
<td><strong>usr:</strong> I would like to visit in town centre please.</td>
<td><strong>sys:</strong> 天安门城楼，簋街小吃和北京欢乐谷都是很不错的地方呢。</td>
</tr>
<tr>
<td></td>
<td>Tiananmen, Gui Street, and Beijing Happy Valley are very nice places.</td>
</tr>
<tr>
<td><strong>usr:</strong> Can you find an Indian restaurant for me that is also in the town centre?</td>
<td><strong>usr:</strong> 我喜欢北京欢乐谷，你知道这个景点周边的酒店都是什么吗？</td>
</tr>
<tr>
<td></td>
<td>I like Beijing Happy Valley. What hotels are around this attraction?</td>
</tr>
<tr>
<td><strong>Schema</strong></td>
<td><strong>sys:</strong> 那可多了，有A酒店, B酒店, C酒店。</td>
</tr>
<tr>
<td><strong>usr:</strong> I want a hotel in San Diego and I want to check out on Thursday next week.</td>
<td><strong>usr:</strong> 太好了，我正打算在景点附近找个酒店住宿呢，知道哪家评分是4分以上，提供叫醒服务的不？</td>
</tr>
<tr>
<td></td>
<td>There are many, such as hotel A, hotel B, and hotel C.</td>
</tr>
<tr>
<td></td>
<td>Great! I am planning to find a hotel to stay near the attraction. Which one has a rating of 4 or higher and offers wake-up call service?</td>
</tr>
</tbody>
</table>

CrossWOZ: A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

<table>
<thead>
<tr>
<th>Type</th>
<th>Single-domain goal</th>
<th>Multi-domain goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
<td>DSTC2</td>
<td>WOZ 2.0</td>
</tr>
<tr>
<td>Language</td>
<td>EN</td>
<td>EN</td>
</tr>
<tr>
<td>Speakers</td>
<td>H2M</td>
<td>H2H</td>
</tr>
<tr>
<td># Domains</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td># Dialogues</td>
<td>1,612</td>
<td>600</td>
</tr>
<tr>
<td># Turns</td>
<td>23,354</td>
<td>4,472</td>
</tr>
<tr>
<td>Avg. domains</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Avg. turns</td>
<td>14.5</td>
<td>7.5</td>
</tr>
<tr>
<td># Slots</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td># Values</td>
<td>212</td>
<td>99</td>
</tr>
</tbody>
</table>

Task-Oriented Dialogue

Initial user state (=user goal)

id=1(Attraction): fee=free, name=?, nearby hotels=？
id=2(Hotel): name= near (id=1),
          wake-up call=yes, rating=?
id=3(Taxi): from=(id=1), to=(id=2),
          car type=? plate number=？

Final user state

id=1 (Attraction): name=Tiananmen Square,
                 fee=free, nearby hotels={[Beijing Capital
Hotel, Guidu Hotel Beijing]}

id=2 (Hotel): name=Beijing Capital Hotel,
             wake-up call=yes, rating=4.6

id=3 (Taxi): from=Tiananmen Square,
             to=Beijing Capital Hotel,
             car type=#CX, plate number=#CP


• 自然語言理解
  Natural Language Understanding (NLU)

• 對話管理
  Dialog Management (DM)

• 自然語言生成
  Natural Language Generation (NLG)

Question Answering and Dialogue Systems

Question Answering (QA)

BERT for Question Answering

Source: Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/

Description: Fine tune pretrained BERT from HuggingFace Transformers on SQuAD.

Introduction

This demonstration uses SQuAD (Stanford Question-Answering Dataset). In SQuAD, an input consists of a question, and a paragraph for context. The goal is to find the span of text in the paragraph that answers the question. We evaluate our performance on this data with the "Exact Match" metric, which measures the percentage of predictions that exactly match any one of the ground-truth answers.

We fine-tune a BERT model to perform this task as follows:

1. Feed the context and the question as inputs to BERT.
2. Take two vectors S and T with dimensions equal to that of hidden states in BERT.
3. Compute the probability of each token being the start and end of the answer span. The probability of a token being the start of the answer is given by a dot product between S and the representatio of the token in the last layer of BERT, followed by a softmax over all tokens. The probability of a token being the end of the answer is compute similarly with the vector T.
4. Fine-tune BERT and learn S and T along the way.

References:

- BERT
- SQuAD

https://tinyurl.com/aintpupython101
Summary

• Artificial Intelligence
• Question Answering
• Dialogue Systems
• Task Oriented Dialogue System
References

- Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/
- Olivier Grisel (2020), Transformers (BERT fine-tuning): Joint Intent Classification and Slot Filling, https://m2dsupsdlclass.github.io/lectures-labs/
  https://github.com/Apress/text-analytics-w-python-2e
  https://www.oreilly.com/library/view/applied-text-analysis/9781491963036/
Q & A

人工智慧任務導向對話系統
（AI Task-Oriented Dialogue System）

中國文化大學應用數學系專題演講
Host: Prof. Chia-Hui Liu
Department of Applied Mathematics, Chinese Culture University
Time: 13:10-15:00, Nov 27, 2020 (Friday)
Place: 理學院會議室 (大義館610), CCU
Address: 55, Hwa-Kang Rd., Yang-Ming-Shan, Taipei, Taiwan

Min-Yuh Day
戴敏育
Associate Professor
副教授
Institute of Information Management, National Taipei University
國立臺北大學 資訊管理研究所

https://web.ntpu.edu.tw/~myday
2020-11-27