

# Deep Learning for Conversational AI

Pei-Hao Su, Nikola Mrkšić, Iñigo Casanueva, Ivan Vulić

POLYAI



NAACL 2018 Tutorial

# Deep Learning for Conversational AI



Eddy



Nikola



Iñigo



Ivan

Tutorial slides are available here:

[poly-ai.com/naacl18](http://poly-ai.com/naacl18)

[#naacl2018](https://twitter.com/naacl2018)

[www.poly-ai.com](http://www.poly-ai.com)

# Conversational AI in Fiction



# Virtual Assistants are the Next Platform



March, 1989.  
Tim Berners-Lee  
invented the World  
Wide Web.



January, 2007  
Apple released the 1st  
iPhone, giving birth to  
the smartphone era.



March, 1876.  
Alexander Graham Bell  
invented the telephone.



2011 - now  
The era of virtual  
personal assistants.

# Virtual Assistants are the Next Platform

## Voice and Language as the Command



# Virtual Assistants are the Next Platform

The Amazon Echo is the top selling item on Amazon

20% of US homes *already own a smart speaker*

- ...and this number will quadruple in the next two years
- 53% of smart speaker owners bought their device last Christmas

Virtual assistants are becoming the main entry point for all actions and services available across all smart devices

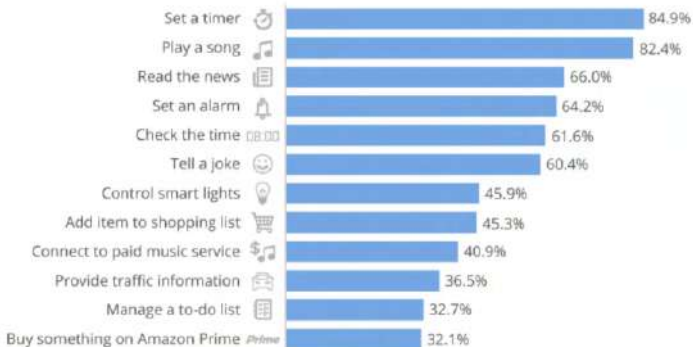
- Instead of different interfaces for smartphones, watches and homes, virtual assistants allow users to access them in the same way – using their voice

# Conversational AI in 2018

## Tech Chart of the Day

### How People Use Amazon's Virtual Assistant

% of Amazon Echo users who have asked the device to do the following at least once



Based on a survey of 180 Amazon Echo users conducted in 2016

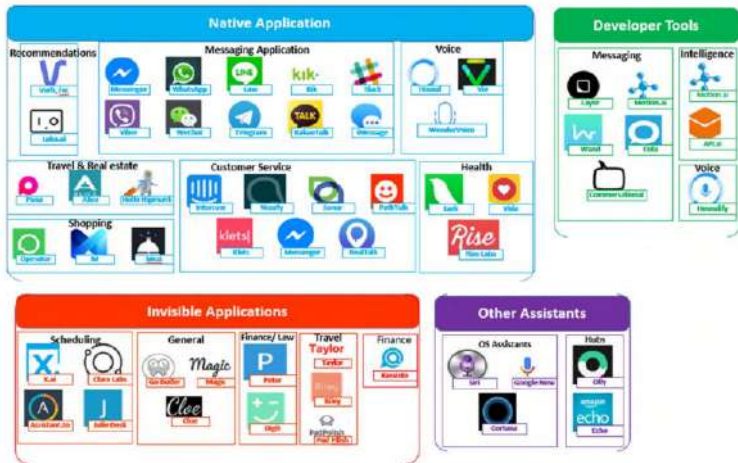
BUSINESS INSIDER

Sources: Experian, Creative Strategies



statista

# Conversational AI = Conversational UI?



**Metis**  
Strategy

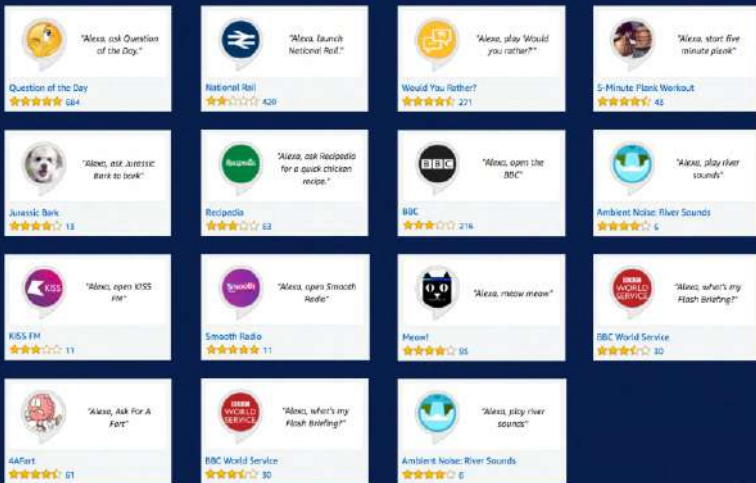
<http://www.larryliu.me/ipa>



# Conversational AI in 2018

Are we there yet?

- Top Alexa skills are still simple single-turn commands



# Conversational AI in 1966

Joseph Weizenbaum's **ELIZA** (1966)

**ELIZA:** What is bothering you?

**User:** I feel sad all the time

**ELIZA:** Why do you feel sad all the time?

**User:** I don't get along with my parents.

**ELIZA:** Tell me about your parents.

# Conversational AI in 2018

## Applications

- **Bookings:** restaurants, hotels, flights, train tickets, ...
- **E-commerce:** from search to payment
- **Accounting and Finance:** from organising records to cryptocurrency trade
- **Healthcare:** from updating the EMRs to psychotherapy
- **Marketing:** e-mail marketing, social media search, ...
- **Daily life:** setting your alarm clock, reminding about meetings, ...

Are we there yet?

# Tutorial Goals

Give an overview of recent research trends in deep learning for conversational AI

Provide a detailed overview of (importance of) task-based dialogue systems

Analyse most promising research avenues and stress their current limitations

Discuss the importance of data requirements vs. algorithm choices

Present an industry-based perspective on current deep conversational AI

Detect current “make it or break it” challenges in conversational AI

# Tutorial Overview

## Part I: Introduction to Conversational AI (20 mins)

- Two paradigms: task-based vs. chatbot-style systems
- Why task-based conversational AI?
- Overview of the full pipeline: core modules

# Tutorial Overview

## Part II: Language Understanding and Dialogue State Tracking *(40 mins)*

- Rule-based approaches and conventional approaches that separate SLU and DST
- Joint approaches: delexicalisation-based
- Joint approaches: methods based on distributed word representations
- Dealing with ASR errors and language ambiguity
- Language understanding across different domains and different languages

# Tutorial Overview

## Part III: Dialogue Management and Reinforcement Learning (60 mins)

- Rule-based and supervised approaches
- RL-based approaches (with RL theory)
- Current challenges with RL-based approaches: evaluation, sample inefficiency, simulation vs. human experience, reward shaping

# Tutorial Overview

## Part IV: Natural Language/Response Generation and Data Collection *(40 mins)*

- Conventional NLG methods: template-based, plan-based, class-based, phrase-based approaches
- RNN-based NLG: regular and attention-based architectures
- End-to-end dialogue modeling as a conditional response generation task
- Why is collecting data so difficult?
- Two recent paradigms for effective data collection



# Tutorial Overview

## Part V: Applications, Challenges, Conclusions (20 mins)

- Recent trends in the development of task-oriented dialogue systems
- Useful software packages and implementations
- Challenges: Scalability, multiple domains, multiple languages, ...

# Conversational AI in 2018

**Google Duplex:** demo released in May 2018

“The technology is directed towards **completing specific tasks**, such as scheduling certain types of appointments. For such tasks, the system makes the conversational experience as natural as possible, allowing people to speak normally, like they would to another person, without having to adapt to a machine...”

“One of the key research insights was to constrain Duplex to **closed domains**, which are **narrow enough** to explore extensively. Duplex can only carry out natural conversations after being deeply **trained in such domains**. It **cannot carry out general conversations**.”

Source: [ai.googleblog.com](https://ai.googleblog.com)

# Challenges and Problems

- How to deal with **variability** and **flexibility** of language?
- What about other languages: **multilinguality**?
- How to build **robust** and **multi-domain** systems?
- How to get **high-quality in-domain data**?
- How to build **meaning representation**?
- How to integrate **commonsense** and **world knowledge**?
- Do we want **very specialised** or more **adaptable** systems?

# The Two Paradigms: Task-Oriented vs Chatbots

M: Hello, welcome to the Cambridge dialogue system, What kind of food would you like?  
H: Yeah I want to find a restaurant that serves European food.  
M: Hotel du Vin and Bistro is a nice place. It serves European food.  
H: Uh what is the address and phone number?  
M: Hotel du Vin and Bistro is a nice place. Their address is XXX and their phone is YYY.  
H: Thank you goodbye.  
M: Thank you for using this system.

[Su et al, 2016]

A: How old are you ?  
B: I'm 16. Why are you asking ?  
A: I thought you were 12.  
B: What made you think so ?  
A: I don't know what you are talking about.  
B: You don't know what you are saying.

[Li et al, 2016]

# The Two Paradigms

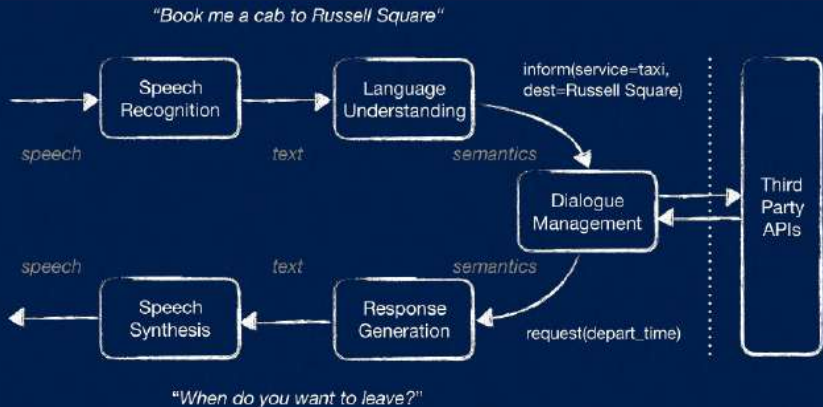
## 1. Task-Oriented Dialogue Systems

- Goal-oriented
- Require precise understanding, it is hard to collect data
- Modular, highly hand-crafted, restricted ability, but meaningful/useful systems

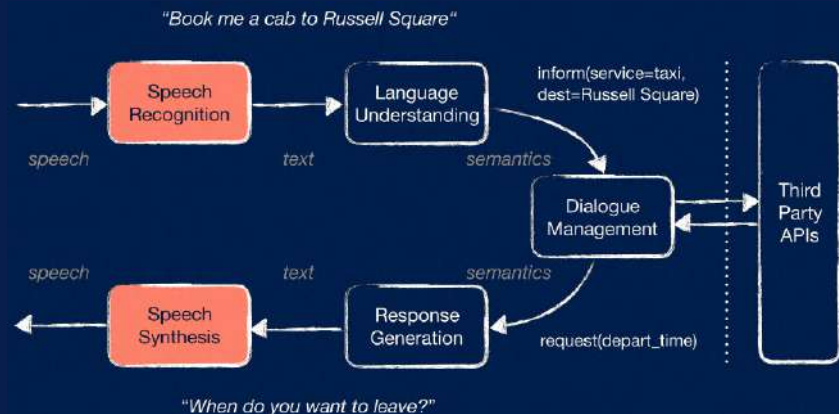
## 2. Chat-Based Conversational Agents

- Chit-chat (no goal)
- Large amounts of data (but probably not helpful so much)
- End-to-end, highly data-driven, but meaningless/inappropriate responses, unreliable systems

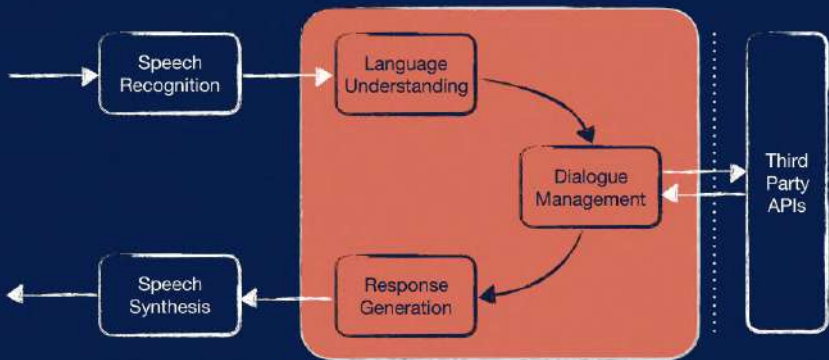
# Task-Oriented Dialogue Systems



# Task-Oriented Dialogue Systems



# Task-Oriented Dialogue Systems





# Dialogue as a Partially Observable Markov Decision Process (POMDP)

[Williams and Young, 2007; Young, 2010; Young et al., 2013]

## Data

- Noisy observations
- Propagating uncertainty and modelling the *belief state*

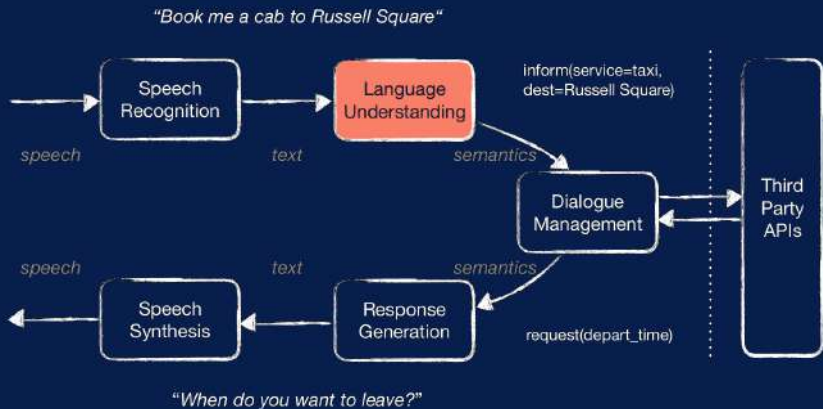
## Model

- Partially observable Markov decision process

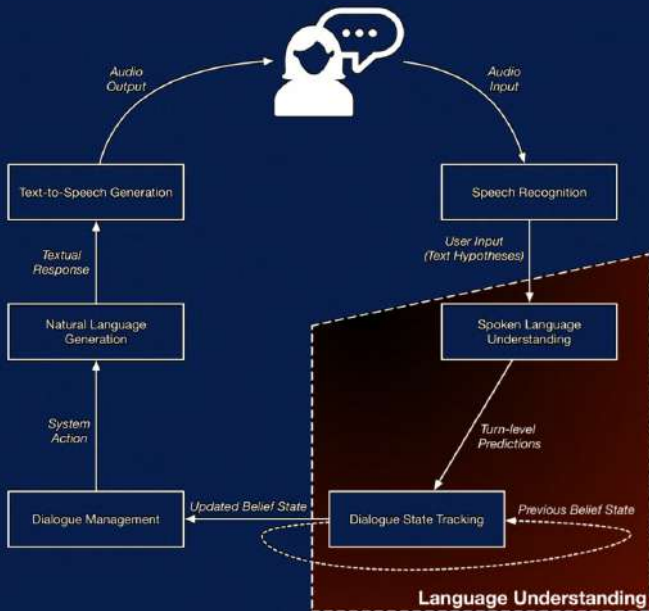
## Predictions

- Optimal system actions in noisy environments

# Natural Language Understanding



# Language Understanding



# Spoken Language Understanding

A lot of “deep work” recently:

- RNNs (encoder-decoder architectures) for slot tagging [Yao et al., 2013; Mesnil et al., 2015; Simmonet et al., 2015; Kurata et al., 2016, Tafforeau et al., 2016]: bidirectional LSTMs, attention-based, multi-task learning
- Joint semantic frame parsing [Hakkani-Tur et al., 2016; Liu and Lane, 2016]: slot filling and intent prediction jointly

but (besides **demanding data/resource requirements**)...

A recent trend has been to combine the two modules into a single *end-to-end learnable* component. Indeed, the top-performing approaches in DSTC2 use both SLU features and n-gram features extracted directly from ASR.

# Spoken Language Understanding



Three main components in the pipelined fashion

# Dialogue Act Classification...

...also known as **intent classification**

Label each dialogue utterance with intent

find me a cheap taiwanese restaurant in oakland

Movies

Find\_movie

Restaurants

Buy\_tickets

Sports

Find\_restaurant

Weather

Book\_table

Music

Find\_lyrics

Deep neural networks can be directly applied to this standard classification problem

[Sarikaya et al., 2011; Tur et al., 2012; Ravuri and Stolcke, 2015; Lee and Dornoncourt, 2016; Khanpour et al., 2016]

# Language Understanding as Slot Filling

## Sequence/Slot Tagging

ATIS UTTERANCE EXAMPLE IOB REPRESENTATION

<b>Sentence</b>	<i>show</i>	<i>flights</i>	<i>from</i>	<i>Boston</i>	<i>To</i>	<i>New</i>	<i>York</i>	<i>today</i>
<b>Slots/Concepts</b>	O	O	O	B-dept	O	B-arr	I-arr	B-date
<b>Named Entity</b>	O	O	O	B-city	O	B-city	I-city	O
<b>Intent</b>	<i>Find Flight</i>							
<b>Domain</b>	<i>Airline Travel</i>							

- Variations of RNN architectures have been used: LSTM-based models, BiLSTMs, other variations... [Mesnil et al., 2015]
- More recently: (attention-based) encoder-decoder architectures instead of basic RNNs [Simmonet et al., 2015; Kurata et al., 2016]
- Multi-task learning formulation: sharing across domains to increase robustness [Jaech et al., 2016; Tafforeau et al., 2016]

# A Note on Semantic Parsing

## SLU and DST resemble the task of **semantic parsing**

Converting natural language into computer-executable formal **meaning representations** for domain-specific applications

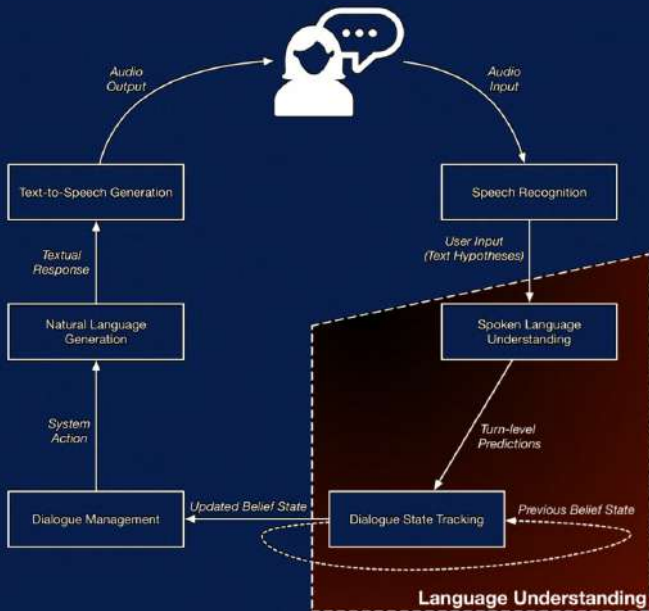
- The dialogue act formalism is considerably simpler
- Do we really need a full meaning representation of each sentence?

Dataset	Example
GEO	how many states border the state with the largest population? answer(A.count(B,(state(B).next_to(B,C).largest(D,(state(C).population(C,D))))),A))
JOBS	what jobs desire a degree but don't use c++? answer(A, (job(A), des_deg(A),+((language(A,C).const(C,'c++')))))
ATIS	what flights from tacoma to orlando on saturday ( _lambda 0e( and( flight0 ) ( _from 0tacoma :e i)( to0 orlando: .ci ) ( _day \$0 saturday: .da ) ) )

[Kate et al., 2010; Liu and Lane, 2016; Hakkani-Tur et al., 2016, Padmakumar et al., 2017, Ravichander et al., 2017]



# Language Understanding



# Dialogue State Tracking Challenge (DSTC)

Challenge	Type	Domain	Data Provider	Main Theme
DSTC1	Human-Machine	Bus Route	CMU	Evaluation Metrics
DSTC2	Human-Machine	Restaurant	Cambridge	User Goal Changes
DSTC3	Human-Machine	Tourist Info	Cambridge	Domain Adaptation
DSTC4	Human-Human	Tourist Info	I2R	Human Conversation
DSTC5	Human-Human	Tourist Info	I2R	Language Adaptation
DSTC6	Human-Human	Restaurant, Twitter	Various	End-to-End Modeling

It is now **Dialogue System Technology Challenge!**

# Domain Ontology

## Domain Ontologies for Task-Oriented Dialogue

Task-oriented systems' ontologies consist of a collection of slots  $s \in S$  (i.e. *food*, *price*, etc.) and their slot values  $V_s$  (*cheap*, *expensive*, etc.).

Good morning, how can I help?

Hi, I'm looking for a Chinese restaurant.

inform ( food = Chinese )

What area would you like?

How about something near Regent Street.

inform ( area = Regent Street )

inform ( food = Chinese )

Szechuan is the only restaurant which serves Chinese food near Regent Street.

What's the address please?

inform ( area = Regent Street )

inform ( food = Chinese )

request ( address )

Szechuan can be found at 15 - 21 Ganton Street.

Awesome, thanks for your help, bye!

simple-act ( goodbye )

Thank you, goodbye!

# Domain Ontology

## Dialogue State Tracking Challenge 2

The **belief state** is a probability distribution over the possible **dialogue states** defined by the **domain ontology**. The downstream **dialogue manager** component uses the distribution to decide on the **next system action**.

DSTC2 Ontology: three *Informable* slots...

**Area:** North, East, South, West, Centre + *dontcare*

**Price:** Cheap, Moderate, Expensive + *dontcare*

**Food:** Indian, Chinese, Thai, Japanese, Greek, French + *84 others*

...and seven (turn-level) *Requestable* slots...

**Food, Area, Price, Address, Phone Number, Postcode**

# Domain Ontology

## A sample from the Cambridge Restaurants domain

INFORMABLE SLOTS: {

PRICE RANGE: [

cheap,  
moderate,  
expensive

],

AREA: [

centre,  
north,  
west,  
south,  
east

],

FOOD: [

Afghan, African, Afternoon Tea, Asian, Australian,  
Austrian, Barbeque, Basque, Belgian, Bistro, Brazilian,  
British, Cantonese, Caribbean, Catalan, Chinese,  
Christmas, Corsican, Creative, Crossover, Cuban,  
Danish, Dutch, English, Eritrean, French, Fusion ...

],

# Dialogue State Tracking Challenge 2: Evaluation

We focus on two key evaluation metrics [Henderson et al., 2014]:

- **Goals** (joint goal accuracy): the proportion of dialogue turns where all the user's search goal constraints were correctly identified;
- **Requests**: similarly, the proportion of dialogue turns where user's requests for information were identified correctly.

# Web-Style Ranking and SLU Combination

[Williams, 2014] uses web-style ranking (lambdaMART algorithm) to rank dialogue states, based on features from three different SLU engines. This was the top entry in the original DSTC2 challenge.

model	Goal		Method		Requested		Requested*	
	Acc.	L2	Acc.	L2	Acc.	L2	Acc.	L2
Best baseline	0.719	0.464	0.897	0.158	0.884	0.196	0.884	0.196
Best DSTC2 result from another team	0.768	<b>0.346</b>	0.940	0.095	<b>0.978</b>	<b>0.035</b>	0.978	0.035
SLU0+1, AB, model comb. (entry0)	0.775	0.758	0.944	0.092	0.954	0.073	0.977	0.037
SLU0+2, AB, model comb. (entry1)	<b>0.784</b>	0.735	<b>0.947</b>	<b>0.087</b>	0.957	0.068	<b>0.980</b>	<b>0.034</b>
SLU0+2, AB, maxent (entry3)	0.771	0.354	<b>0.947</b>	0.093	0.941	0.090	0.979	0.040

# Language Understanding

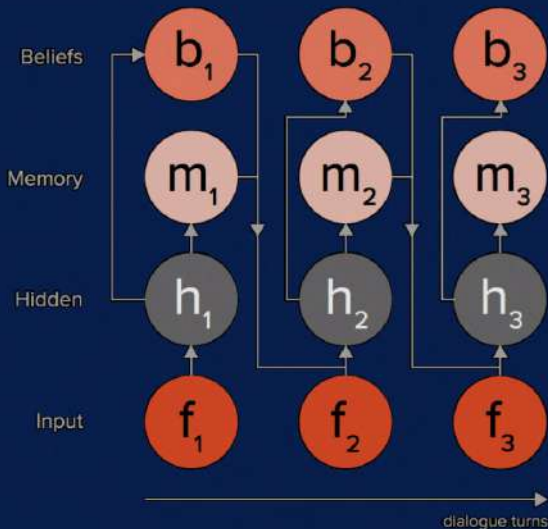
## Joint vs. Separate SLU?

Entry	Features		Joint Goals		Search Method		Requested	
	ASR	SLU	Acc.	L2	Acc.	L2	Acc.	L2
1-best baseline <sup>1</sup>		✓	0.619	0.738	0.879	0.209	0.884	0.196
focus baseline <sup>1</sup>		✓	0.719	0.464	0.867	0.210	0.879	0.206
HWU baseline <sup>2</sup>		✓	0.711	0.466	0.897	0.158	0.884	0.201
team1 entry0		✓	0.601	0.648	0.904	0.155	0.960	0.073
team3 entry0		✓	0.729	0.452	0.878	0.210	0.889	0.188
team4 entry2		✓	<b>0.742</b>	<b>0.387</b>	<b>0.922</b>	<b>0.124</b>	<b>0.957</b>	<b>0.069</b>
team6 entry2		✓	0.718	0.437	0.871	0.210	0.951	0.085
team7 entry4		✓	0.735	0.433	0.910	0.140	0.946	0.089
team8 entry1		✓	0.699	0.498	0.899	0.153	0.939	0.101
team9 entry0		✓	0.499	0.760	0.857	0.229	0.905	0.149
team2 entry2	✓		0.668	0.505	<b>0.944</b>	0.095	0.972	0.043
team4 entry0	✓		<b>0.768</b>	<b>0.346</b>	0.940	<b>0.095</b>	<b>0.978</b>	<b>0.035</b>
team7 entry0	✓		0.750	0.416	0.936	0.105	0.970	0.056
team2 entry1	✓	✓	<b>0.784</b>	0.735	<b>0.947</b>	<b>0.087</b>	0.957	0.068
team2 entry3	✓	✓	0.771	<b>0.354</b>	0.947	0.087	0.941	0.090
team5 entry4	✓	✓	0.695	0.610	0.927	0.147	<b>0.974</b>	<b>0.053</b>



# Word-Based DST Using RNNs

[Henderson et al., 2014]



At each dialogue turn, for each slot in the ontology:

**Beliefs:** a probability distribution over the available slot values (the belief state) for each slot in the ontology.

**Memory:** a continuous vector representing dialogue context.

**Input:** lexical and delexicalised features extracted from the current user utterance, previous system act, belief state and the RNN memory layer.

# Delexicalisation-Based DST

## Delexicalised Features

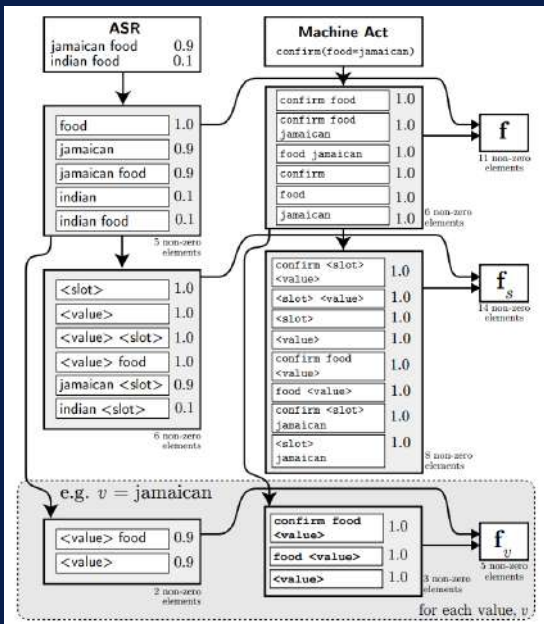
**Delexicalised features:** all occurrences of slot names and/or slot values in an utterance are replaced with **generic tags**

I want Chinese food  
I want cheap price range ] I want VALUE SLOT

The use of delexicalised  $n$ -gram features facilitates:

- Faster learning by facilitating transfer learning across slot values
- Generalisation to unseen slot values (or even entirely new slots)
- Bootstrapping dialogue systems to new domains with limited data

# Delexicalisation-Based DST



# Delexicalisation-Based DST

## Towards Multi-Domain DST

Delexicalised features allow transfer learning between slots

Delexicalisation facilitates **multi-domain** training

- RNN-based DST with **shared initialisation** + **slot-specialised training** [Mrkšić et al., 2015]



# Delexicalisation-Based Models

## Shortcomings

### Delexicalised Features = Exact Matching

Given an arbitrary domain ontology, delexicalisation-based models provide data-efficient language understanding - as long as users use only the actual ontology values to express their search constraints!

**User:** I'm looking for an affordable restaurant  
`inform(price=cheap)`

**System:** How about Thai food?

**User:** Yes please, in central Cambridge  
`inform(price=cheap, food=Thai, area=centre)`

**System:** The House serves cheap Thai food

**User:** Where is it?  
`inform(price=cheap, food=Thai, area=centre);  
request(address)`

**System:** The House is at 106 Regent Street

# Traditional 'Solution': Semantic Dictionaries

## Delexicalised Features = Exact Matching

Delexicalisation-based models allow fast deployment to new dialogue domains, but introduce a complete dependency on **semantic dictionaries**.

**Food=Cheap:** [affordable, budget, low-cost, low-priced, inexpensive, cheaper, economic, ...]

**Rating=High:** [best, high-rated, highly rated, top-rated, cool, chic, popular, trendy, ...]

**Area=Centre:** [center, downtown, central, city centre, midtown, town centre, ...]

# The Neural Belief Tracker (NBT)

## Semantic Dictionaries Re-Introduce the SLU module

Semantic dictionaries can be hand-crafted or learned - but only for simple toy domains such as DSTC2. Moreover, the Amazon Mechanical Turk data collection framework forces the users to use very simple (and unnatural) language, understating the challenge of dealing with linguistic variation.

# The Neural Belief Tracker (NBT)

## Semantic Dictionaries Re-Introduce the SLU module

Semantic dictionaries can be hand-crafted or learned - but only for simple toy domains such as DSTC2. Moreover, the Amazon Mechanical Turk data collection framework forces the users to use very simple (and unnatural) language, understating the challenge of dealing with linguistic variation.

The **Neural Belief Tracker** is a statistical DST model/framework which aims to satisfy the following design goals:

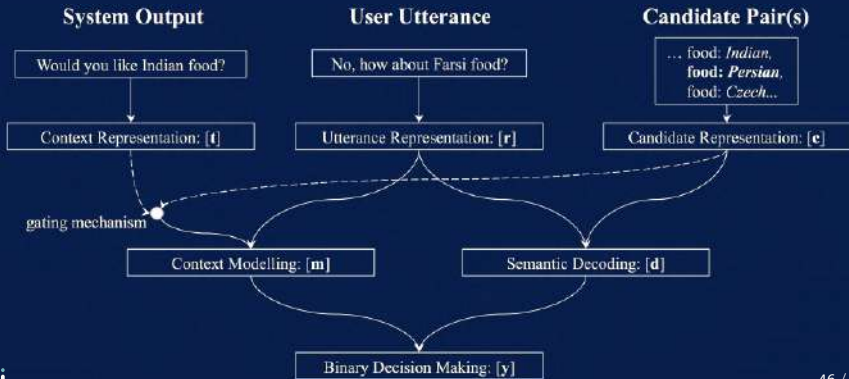
- End-to-end learnable (no SLU modules or semantic dictionaries).
- Generalisation to unseen slot values.
- Capability of leveraging the semantic content of pre-trained word vector spaces without human supervision.



# The Neural Belief Tracker (NBT)

## Representation Learning + Label Embedding + Separate Binary Decisions

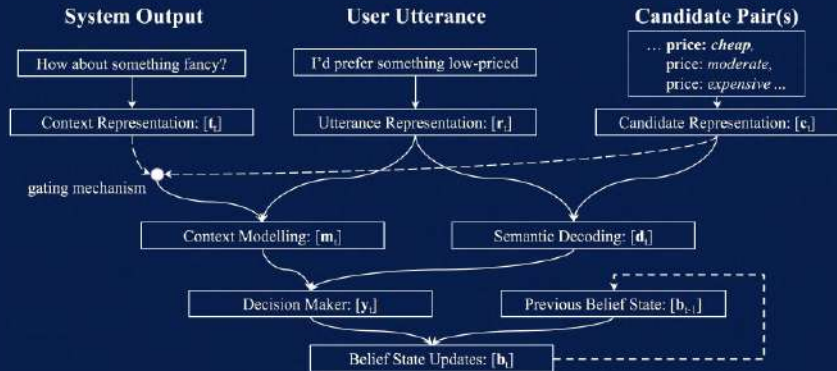
To overcome data sparsity, NBT models use *label embedding* to decompose multi-class classification into many binary ones.



# The Neural Belief Tracker (NBT) v2

## Fully Statistical Belief State Update

Removing the last rule-based component of the NBT framework

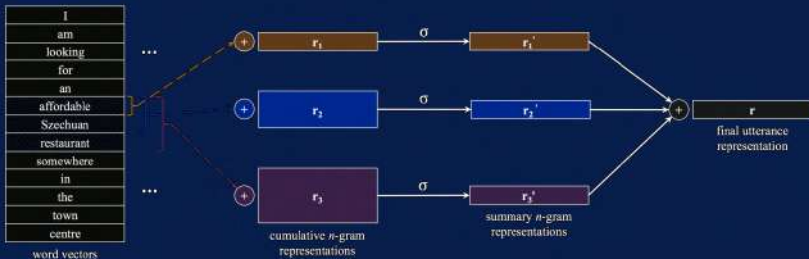


# Representation Learning: NBT-DNN

Let  $u$  represent a user utterance consisting of  $k_u$  words  $u_1, u_2, \dots, u_{k_u}$ .

Each word has an associated **fixed** word vector  $u_1, \dots, u_{k_u}$

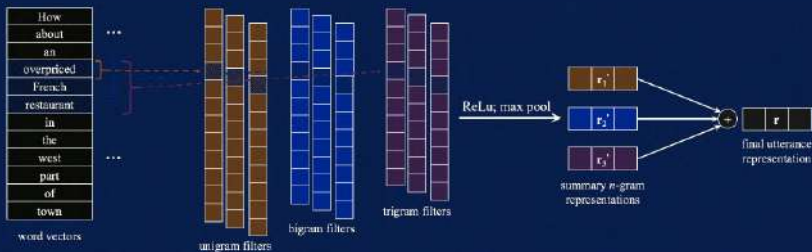
$$\mathbf{v}_i^n = \mathbf{u}_i \oplus \dots \oplus \mathbf{u}_{i+n-1}$$



$$\mathbf{r}_n = \sum_{i=1}^{k_u-n+1} \mathbf{v}_i^n; \quad \mathbf{r}'_n = \sigma(W_n^s \mathbf{r}_n + b_n^s); \quad \mathbf{r} = \mathbf{r}'_1 + \mathbf{r}'_2 + \mathbf{r}'_3$$

# Representation Learning: NBT-CNN

Let  $F_n^s \in R^{L \times nD}$  denote the collection of filters for each value of  $n$ , where  $D = 300$  is the word vector dimensionality. If  $\mathbf{v}_i^n$  denotes the concatenation of  $n$  **fixed** word vectors starting at index  $i$ , let  $\mathbf{m}_n = [\mathbf{v}_1^n; \mathbf{v}_2^n; \dots; \mathbf{v}_{k_u-n+1}^n]$  be the list of  $n$ -grams that convolutional filters of length  $n$  run over.

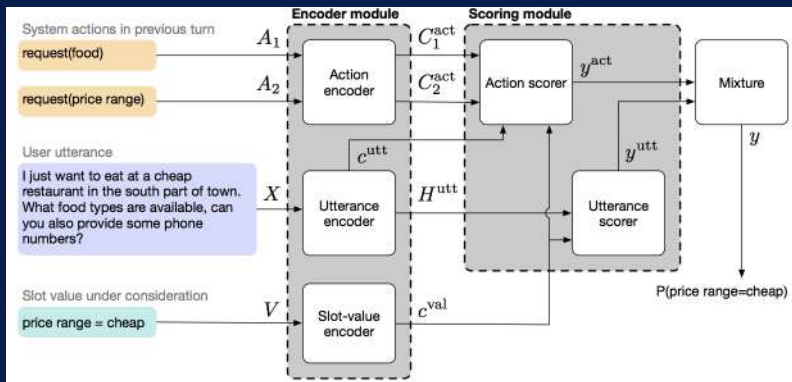


$$R_n = F_n^s \mathbf{m}_n \quad \mathbf{r}'_n = \text{maxpool}(\text{ReLU}(R_n + b_n^s)) \quad \mathbf{r} = \mathbf{r}'_1 + \mathbf{r}'_2 + \mathbf{r}'_3$$

# Global-Locally Self-Attentive DST (GLAD)

[Zhong et al., ACL 2018]

The same idea as NBT, but a more advanced encoder



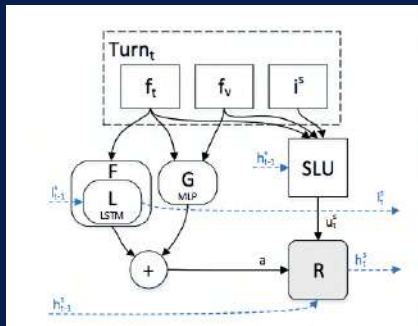
**Global** modules to share parameters between estimators for each slot and **local** modules to learn slot-specific feature representations.

# DST as Machine Reading / Reasoning

[Perez and Liu, 2017] use Memory Networks and pose DST as a machine reading / multi-hop reasoning problem.

Locutor	Utterance	Hop 1	Hop 2	Hop 3	Hop 4	Hop 5
Cust	Im looking for a cheap restaurant that serves chinese food	0.00	0.18	0.11	0.04	0.00
Agent	What part of town do you have in mind	0.33	0.30	0.00	0.00	0.00
Cust	I dont care	0.00	0.00	0.17	0.37	1.00
Agent	Rice house serves chinese food in the cheap price range	0.01	0.00	0.00	0.00	0.00
Cust	What is the address and telephone number	0.58	0.09	0.01	0.00	0.00
Agent	Sure rice house is on mill road city centre	0.03	0.00	0.00	0.00	0.00
Cust	Phone number	0.00	0.00	0.00	0.00	0.00
Agent	The phone number of rice house is 765-239-09	0.02	0.01	0.00	0.00	0.00
Cust	Thank you good bye	0.02	0.42	0.71	0.59	0.00
What is the area ? Answer: dontcare						

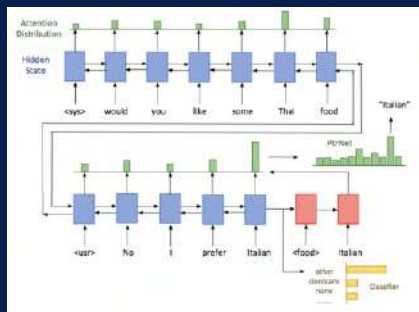
# Hybrid Dialog State Tracker with ASR Features



[Vodolán et al., 2017] propose a *Hybrid Dialog State Tracker*, achieving the current state-of-the-art performance for DSTC2 (0.796).

Their model uses a separate (learned) SLU decoder coupled with a rule-based system to achieve the best DSTC2 performance to date.

# An End-to-end Approach for Handling Unknown Slot Values in Dialogue State Tracking



[Xu and Hu, 2017] use *Pointer Networks* to address the problem of dealing with *unknown*, rather than *unseen* values.

This is an important direction for future work, as most presented models rely on a canonical representation of the domain ontology that is *predefined* by the dialogue system designers.



# Are Word Embeddings Useful for NLU?

# Input Word Vectors Matter

## Semantic Specialisation

It is essential that deployed word vectors are specialized for true semantic similarity [Mrkšić et al.; NAACL 2016, TACL 2017]

**User:** I'm looking for a cheaper restaurant  
inform(price=expensive)

**System:** What kind of food?

**User:** English, in eastern Cambridge  
inform(price=expensive, food=Spanish, area=east)

**System:** The Green Man is the best choice

**User:** Where is it?  
inform(price=expensive, food=Spanish, area=east);  
request(address)

**System:** The Green Man is at 59 High St, Grantchester

# Input Word Vectors Matter

## Semantic Specialisation

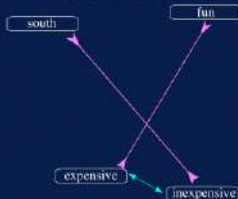
Take a mini-batch of ATTRACT and REPEL pairs...

For each pair, find two *pseudo-negative* examples... ..and fine-tune the vectors so that ATTRACT pairs are closest...

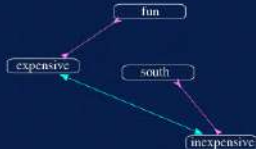
ATTRACT



REPEL



...and REPEL pairs furthest away from each other



# Semantic Specialisation

## Semantically-Specialising Word Vector Spaces

Injecting semantic constraints (such as synonymy and antonymy) into pre-trained word vectors can mitigate some of the negative aspects of **the distributional hypothesis**.

Word	east	expensive	British
Before	west	pricey	American
	north	cheaper	Australian
	south	costly	Britain
	southeast	overpriced	European
	northeast	inexpensive	England
After	eastward	costly	Brits
	eastern	pricy	London
	easterly	overpriced	BBC
	-	pricey	UK
	-	afford	Britain

Table: Nearest neighbours before and after semantic specialisation

# The Importance of Semantic Specialisation

Three different word vector collections: **1)** 'random' word vectors initialised using the XAVIER initialisation; **2)** distributional GloVe vectors; and **3)** *semantically specialised* Paragram-SL999 vectors.

Word Vectors	DSTC2		WOZ 2.0	
	Goals	Requests	Goals	Requests
xavier	64.2	81.2	81.2	90.7
GloVe	69.0*	96.4*	80.1	91.4
Paragram-SL999	<b>73.4*</b>	<b>96.5*</b>	<b>84.2*</b>	<b>91.6</b>

Recent developments: specialising **the full vocabulary** yields additional gains

[Vulić et al., NAACL 2018, ACL 2018]

# Towards Multilingual DST

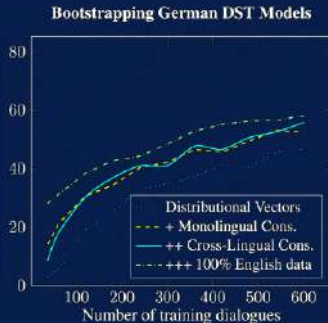
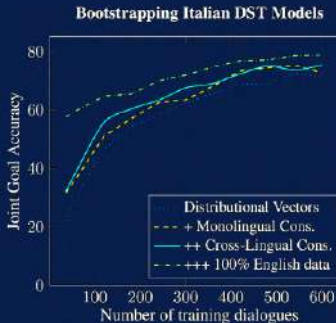
We're still (too) obsessed with English!



# Bootstrapping DST for Resource-Poor Languages

## Ontology Grounding: Multilingual DST Models

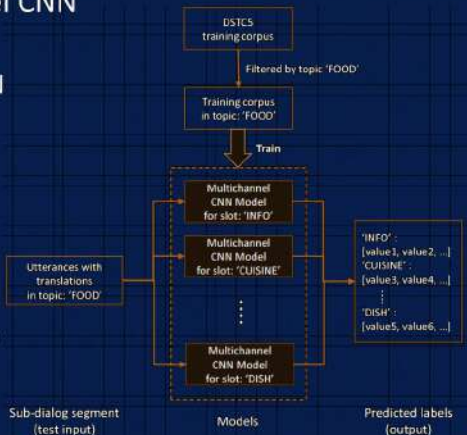
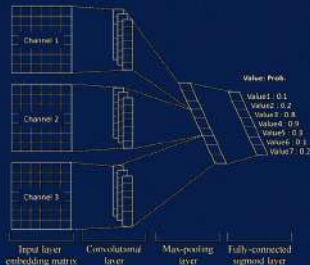
The concepts expressed in the dialogue domain ontology are language-agnostic: 'labels' persist across languages. Given training data for two (or more) languages, and a cross-lingual vector space of high quality, we can train a multilingual DST model. [Mrkšić, Vulić et al., 2017]



# Bootstrapping DST for Resource-Poor Languages

Multichannel DST [Shi et al., 2017]

- Training a multichannel CNN for each slot
  - ▣ Chinese character CNN
  - ▣ Chinese word CNN
  - ▣ English word CNN

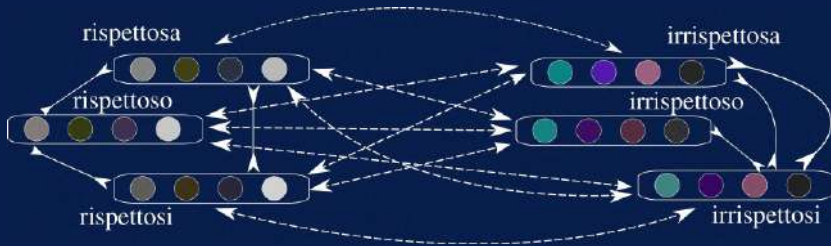




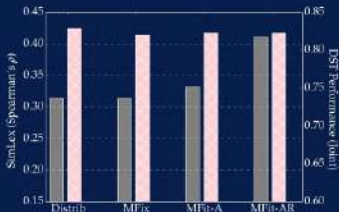
# DST Beyond English: Additional Challenges

## Using Different Sources of Supervision

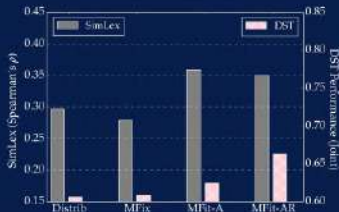
Morphological phenomena provide an inexpensive source of supervision. Consider both *inflectional* and *derivational* forms.



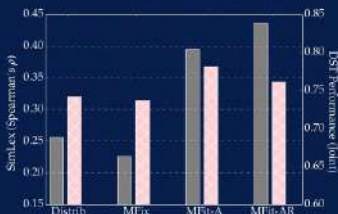
# DST Beyond English with Morph-fitting



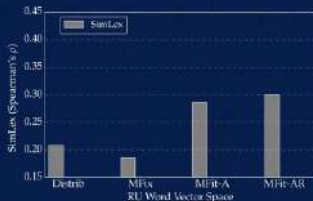
(a) English



(b) German

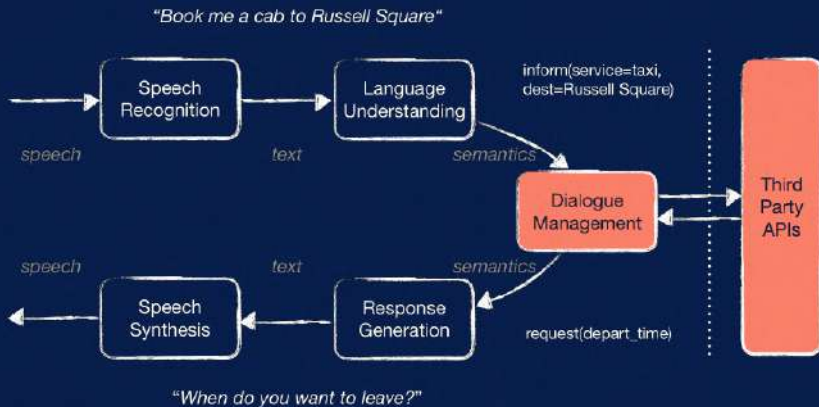


(c) Italian



(d) Russian

# Dialogue Management



# Dialogue Management Approaches

## Rule-Based:

- Huge hand-crafting effort
- Non-adaptable and non-scalable
- But this is what works right now

## Supervised:

- Learn to “mimic” the answers of a corpus
- Assumes optimal human behaviour
- Does not do long-term planning

## Reinforcement Learning (RL)

# Reinforcement Learning

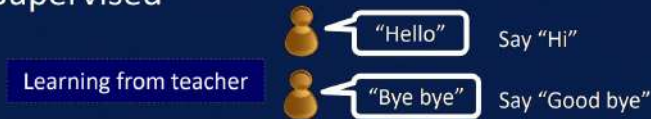
RL is a general-purpose framework for **decision making**

- RL concerns an **agent** with the capacity to **act**
- Each **action** influences the observations obtained from the **environment**
- The agent updates its internal **state** based on these **observations**
- Success is measured by a **reward** signal
- Goal: **select actions to maximize future reward**

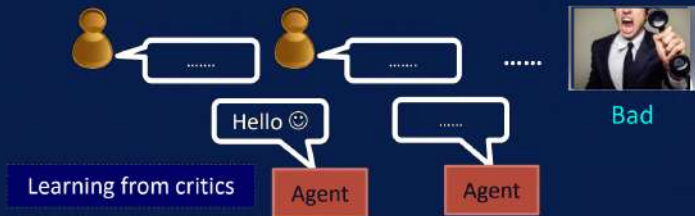


# Reinforcement vs Supervised Learning

## Supervised



## Reinforcement



# Dialogue Management

## Based on RL

- Learns through interaction in order to maximise a future reward
- Learns in the actual dialogue environment
- Adapts to new environments/users/situations
- Requires less annotation
- Slow and expensive learning
- Difficult to reuse data

# Dialogue Management

## Based on RL

### Policy-based RL

- Search directly for optimal policy  $\pi^*$  (i.e., policy achieving maximum future reward)

### Value-based RL

- Estimate the optimal value function  $Q^*(s, a)$  (i.e., maximum value achievable under any policy)

### Model-based RL

- Build a model of the environment
- Plan (e.g., by look-ahead) using the model



# Value-Based Models

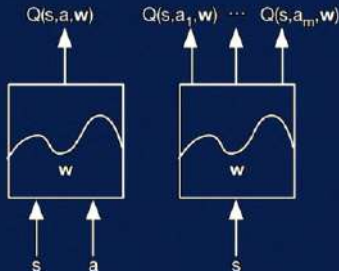
## Q-Networks

**Q-networks** represent value functions with weights  $w$ :

$$Q(s, a, w) \approx Q^*(s, a)$$

[Sutton, 1998]      Deep Q-networks [Mnih et al., 2013]

- They generalize from seen states to unseen states
- They update parameter  $w$  for function approximation



# Value-Based DM Models

Model Q-Value function:

- $Q^\pi(b, a) = E\{R_t | b_t = b, a_t = a, \pi\}$
- $R_t = \sum_{\tau=t}^{T-1} \gamma^{(\tau-t)} r_\tau$ , where  $t$  is the current timestep

The optimal policy can be obtained by greedily taking the action which maximises  $Q$

- $\pi(b) = \arg \max_{a \in A} Q^\pi(b, a)$

Value function can be represented by a deep Q-network with weights  $w$

# Policy-Based DM Methods

## Deep RL

$\pi(a|\mathbf{b})$ ; Approximate with DNN: Scalable, flexible structure

Directly optimise the (stochastic) policy:

$$\nabla_{\Theta} J(\Theta) = \mathbb{E}_{\tau \sim \pi_{\Theta}} \left[ \sum_{t=0}^{T-1} \nabla_{\Theta} \log \pi(a_t | \mathbf{b}_t) \left( \sum_{t'=t}^{T-1} r_{t'} - b(\mathbf{b}_t) \right) \right]$$

Actor-Critics:

$$\begin{aligned} \nabla_{\Theta} J(\Theta) &= \nabla_{\Theta} \mathbb{E}[R | \pi_{\Theta}] = \mathbb{E}[\nabla_{\Theta} \log \pi_{\Theta}(a | \mathbf{b}) A_w(\mathbf{b}, a)] \\ A_w(\mathbf{b}, a) &= Q(\mathbf{b}, a) - V(\mathbf{b}) \end{aligned}$$

**Issues:** sample inefficiency, cold start, slow learning...

# Dialogue Evaluation = Reward for RL

Dialogue is a special RL task

- Human is involved in **interaction** and **rating (evaluation)** of dialogue
- Fully **human-in-the-loop** framework

**Rating:** correctness, appropriateness, adequacy

Expert rating	high quality, <b>high</b> cost
User rating	unreliable quality, <b>medium</b> cost
Objective rating	Check desired aspects, <b>low</b> cost

# Reinforcement Learning Signal

So what is the reward?

In theory: **user satisfaction**

In reality: how do we **measure the satisfaction**?

- Ask the user for feedback after each dialogue → Very inefficient/annoying
- Hand-craft a “user satisfaction” estimator (e.g. success/length trade-off)  
→ We usually need to know the user goal to succeed
- Train a “user satisfaction” estimator using user feedback → We ask for user feedback only when we are uncertain about it [Su et al., 2016]

In sum: reward modeling is still an **open question**

# Dialogue Management

## Evaluation and Reward

How NOT to evaluate your dialogue system

[Liu et al., 2017]

- There is *no standardised metric* and *no standardised benchmark*
- Turn-level or dialogue-level evaluation?
- Other metrics? BLEU, METEOR, ROUGE? Embedding-based metrics?
- Human satisfaction and “objective metrics” do not correlate well
- Automatic evaluation metrics can be biased

# Dialogue Management

## Reinforcement Learning Signal in Task Oriented Settings

Typical Reward Function:

- per turn penalty -1
- Large reward at completion if **successful**
- Typically requires **prior knowledge** of the task
  - ✓ Simulated user
  - ✗ Paid users (Amazon Mechanical Turk)
  - ✗ Real users



The user simulator is usually required for dialogue system training before deployment

# Dialogue Management

## Reinforcement Learning Signal

How to learn policy from real users?

- Infer success (reward) directly from dialogues
  - Train a reward estimator from data (Su et al. 2015)



- User rating
  - Noisy
  - Difficult/Costly to obtain



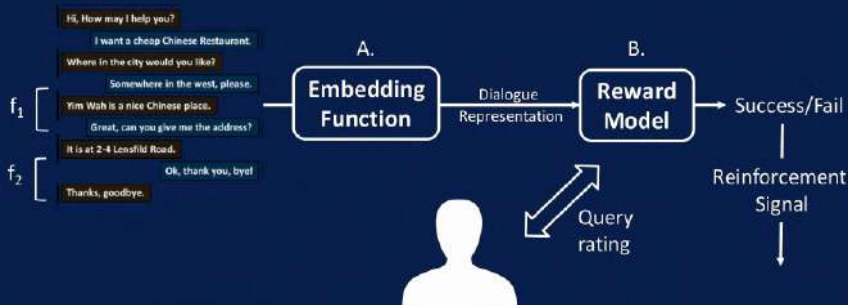
- Robust user rating model (Su et al. ACL, 2016)
  - Noisy → Gaussian Process with uncertainty
  - Difficult/Costly → Active Learning



# Dialogue Management

## Reinforcement Learning Signal

Reward modeling based on users' **binary success rating**

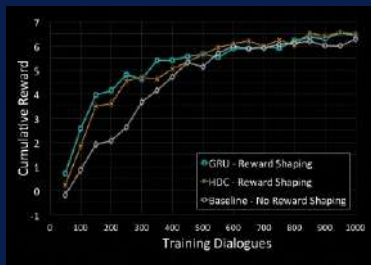
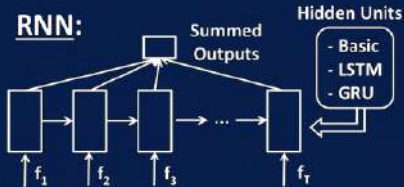


# Reward Shaping for DM

[El Asri et al., 2014; Su et al., 2015]

Extra reward signal  $F$  in addition to environmental reward  $R$ , making the system learn from the composite signal  $R + F$

Turn-level rewards via RNNs



# RL-Based DM: Quick Recap

Continuous state Markov Decision Process (MDP) composed by:

- Multivariate belief state space  $B$
- Finite set of (discrete) actions  $A$
- Reward function  $R(b, a) = r_t$  where  $a \in A, b \in B$

One of the critical challenges: **sample-efficiency**

Solutions: Sample-efficient actor-critic

- Off-policy learning with experience replay
- Better Gradient Update

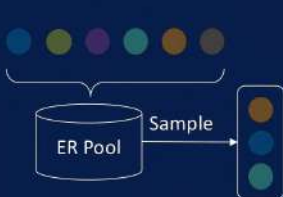
# Sample efficient RL

## Experience Replay (ACER, TRACER, eNACER)

Make efficient use of past data

[Su et al., 2017; Wang et al., 2017; Weisz et al., 2018]

### ➤ On-policy learning



Behavior policy  $\mu(a|x)$



(Munos et al. 2017)

### ➤ Off-policy learning with Experience Replay (ER)

- Correct sampling bias via Importance sampling ratio  $\frac{\pi(a_s|x_s)}{\mu(a_s|x_s)}$

# Dialogue Management

## Training Approaches

### Supervised Learning (SL)

- Wizard-of-Oz (WoZ) methods to collect (high-quality?) data
- Easy to get to reasonable performance
- No planning, not optimal in deploying environments
- Low coverage

### Reinforcement Learning (RL)

- Learn policy online with real or simulated users
- Poor performance in early training stages

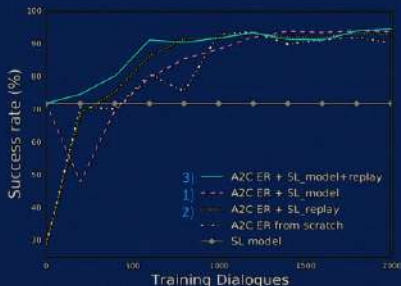
Combine the best of both worlds using one single model

[Rieser and Lemon, 2016; Henderson, 2008; Fatemi, 2016; Zhao, 2016; Williams, 2017]

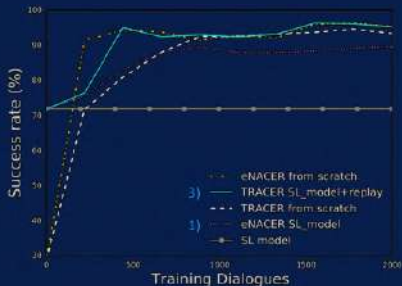
# Sample-Efficient RL + Supervised Data

## Some Empirical Results

[Su et al., 2017]



(a) Learning for A2C ER with demonstration data.



(b) Learning for TRACER and eNACER with demonstration data.

- 1) SL model: Pre-training with SL data
- 2) SL replay: Supervised replay pool  $P_{sup}$  for RL
- 3) SL model+replay: a + b

# User Simulation

## Why don't use a corpus?

- Different policy  $\leftarrow$  different trajectories: Corpus based evaluation is not possible
- Corpus coverage biased / not wide enough

## Replace the real user in the pipeline by a model that replicates the user behaviour

- Cheap / reusable method to train (and test) statistical dialogue systems
- Useful for research, hyperparameter tuning, etc.
- Necessary evil

# Different simulation approaches

## Model:

- **Rule-based:** Manually crafted by experts (synthetic data) [Schatzman et al., 2007; Keizer et al., 2010]
- **Learning-based:** Trained from a corpus of dialogues (data augmentation) [Pietquin and Dutoit, 2006; El Asri et al., 2016; Crook and Martin, 2017]

## Output level:

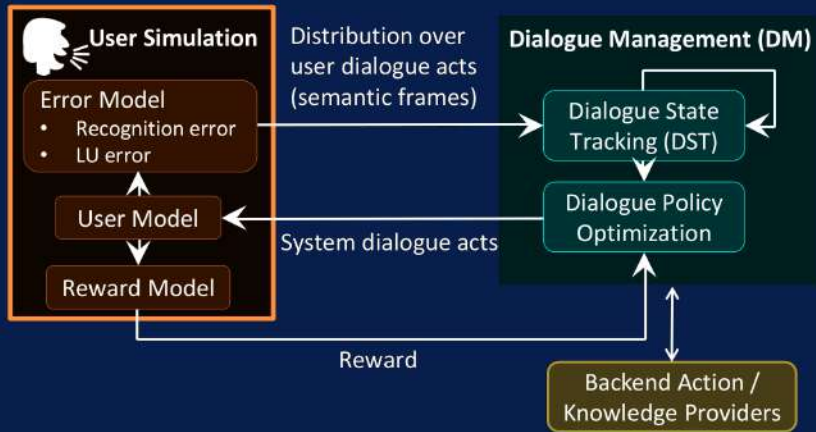
- Semantic level (dialogue act)
- Text level (natural language)

## Error simulation:

- No error
- Language understanding error [Schatzman et al., 2007; Thomson et al., 2010]
- ASR error [Pietquin et al., 2002]



# Training with User Simulator



# User Simulation (Rule-Based)

Trained experts construct and fine-tune the simulation process

Pros:

- No data needed
- Full control of user behaviour

Cons:

- Very costly
- Not enough coverage / variability

# Agenda-Based user simulation

[Schatzman et al., 2007; Keizer et al., 2010]

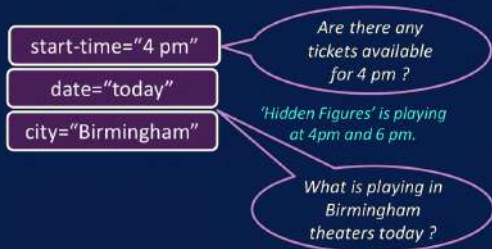
First, generate a user goal.

The user goal contains:

Dialog act

Inform slots

Request slots



```
{  
  "request_slots": {  
    "ticket": "UNK",  
    "theater": "UNK"  
  },  
  "diaact": "request",  
  "inform_slots": {  
    "city": "birmingham",  
    "numberofpeople": "2",  
    "state": "al",  
    "starttime": "4 pm",  
    "date": "today",  
    "moviename": "deadpool"  
  }  
}
```

# User Simulation (Model/Learning-Based)

Transform a static corpus of dialogues into a dynamic tool

Previously:

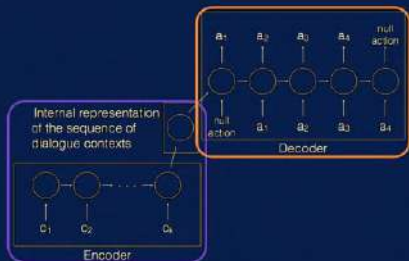
- bigram models [Eckert et al., 1997; Levin et al., 2000]
- graph-based [Scheffler and Young, 2000]
- HMMs [Georgila et al., 2005; Cuayahuitl et al., 2005]
- ...

Today: **neural models**

- Seq2Seq user simulation [El Asri et al., 2016; Crook and Marin, 2017; Kreyssig et al. 2018]

# Seq2Seq User Simulation

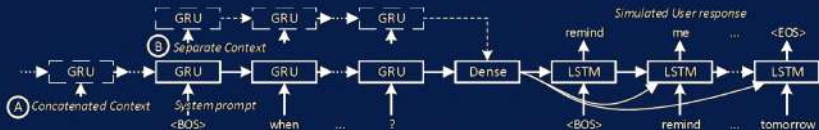
[El Asri et al., 2016]



- **Input:**  $c_i$  encodes contextual features (e.g., previous system action, consistency between user goals and system-provided values)
- **Output:** a dialogue act sequence from the user

# Seq2Seq User Simulation

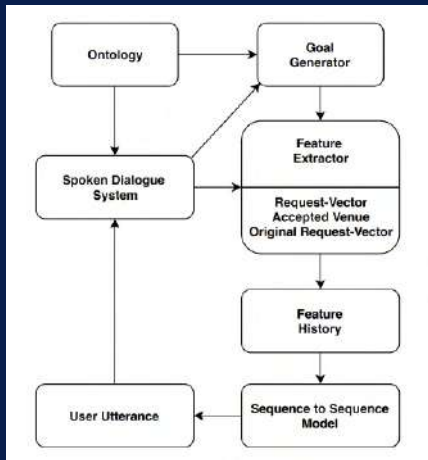
[Crook and Marin, 2017]



- Text based simulation
- No labeled data
- Trained on human-to-machine conversations

# Seq2Seq User Simulation

[Kreyssig et al. 2018]



- **Input:** Contextual features
- **Output:** Natural language
- User goal and goal changes explicitly modelled → Can be used to train a RL-based policy

# Evaluating User Simulation

## Direct methods

- Measure the similarity between the outputs of the simulator and a real user in a test set
- Simulators trained on the same or similar metrics
- Variability is penalised, even if its consistent

## Indirect methods

- Cross-model evaluation
- Training with simulated user / evaluating with real users



# Indirect evaluation

[ Kreyssig et al. 2018]

Training Simulator	Human Evaluation	
	Rew.	Suc.
NUS - $\mathcal{N}_1$	13.4	91.8
NUS - $\mathcal{N}_2$	13.8	93.4
ABUS - $\mathcal{A}_1$	13.3	90.0
ABUS - $\mathcal{A}_2$	13.1	88.5

- RL policy trained with simulated user and tested with real users
- Systems trained with a (good) simulated user perform well with real users
- System performance evaluation using simulated users correlates with performance evaluation with real users

# Reusing data (transfer learning)

## Domain transfer

- Use data collected in one domain to train / bootstrap a dialogue manager in another domain
- [Gašić et al., 2013, 2015; Wang et al., 2015; Papangelis et al., 2017]

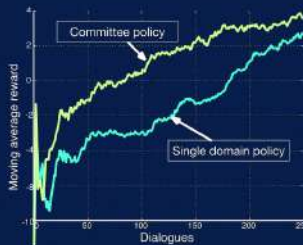
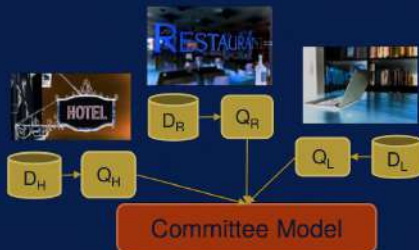
## User adaptation

- Use data collected from a specific user to build an adapted dialogue manager for another user
- [Casanueva et al., 2015; Genevay et al., 2016; Laroche et al., 2017]

# Domain Adaptation

[Gašić et al., 2015]

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

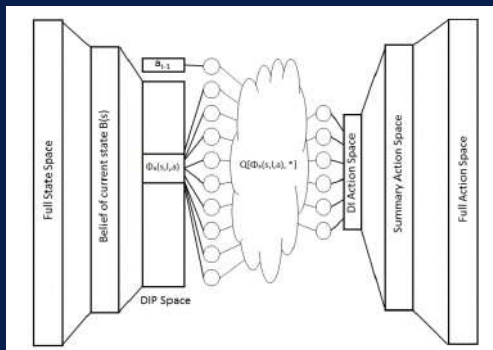


The policy for a new domain can be boosted by the committee policy

# Domain Adaptation

## Domain independent parametrisation (DIP)

[Wang et al., 2015; Papangelis et al., 2017]



- Map the domain specific dialogue state space into a domain independent space

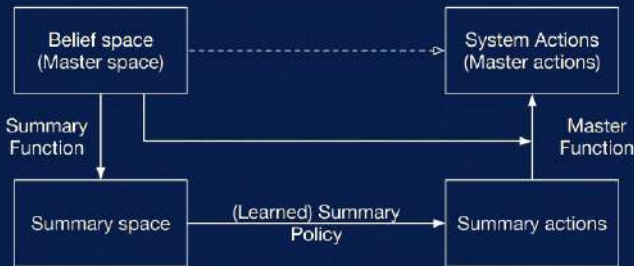
# Battling Data Inefficiency

**More structure → Less data needed**

- Summary actions
- Action masks
- Hierarchical structures

# Summary Actions

Action clustering to reduce the action set  $A$



Summary mapping function:  $a_m = sm(a, b)$

Summary action: `confirm(food)` → Master action: `confirm(food=Asian)`

Belief state can also be summarised

# Action Masks

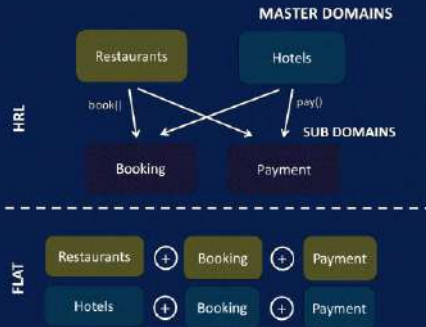
heuristics to reduce the action set  $A$  in each turn

- Mask `confirm()` actions if there is no belief assigned to particular slots
- Mask `inform()` actions if there is not enough evidence of user's goal

Action masks improve efficiency

[Thomson 2010; Williams et al., 2017; Weisz, 2017]

# Dialogue Management Based on Hierarchical RL



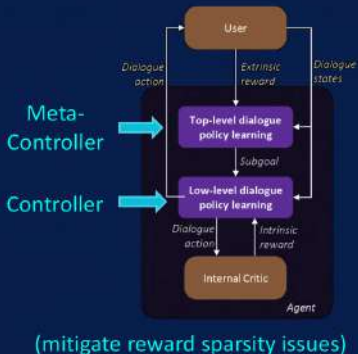
## Hierarchical Policy

Control architectures and learning algorithms specifying a hierarchy of tasks and reusing parts of the space across sub-tasks

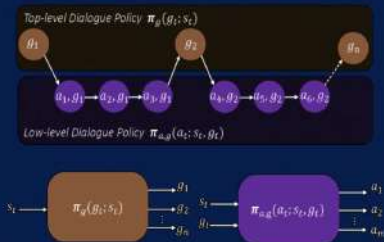


# Dialogue Management Based on Hierarchical RL

[[Cuayáhuitl, 2009; Peng et al., 2017; Budzaniowski 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  $g_t$



## Temporal abstraction

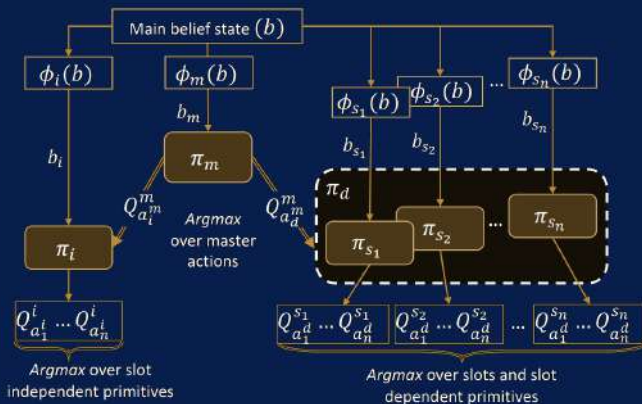
Decisions at the given level are not required at each step but can call temporally extended sub-tasks with their own policies

# Dialogue Management Based on Feudal RL

## Space abstraction

Decisions in each turn are divided into several steps, each step using a subset of the dialogue state space

[Casanueva et al., 2018]



# Revisiting Problems of RL-Based DM

Training from scratch with real users is very expensive and slow

- **Simulated users:** necessary evil
- Pretraining with supervised learning (but we need the data!)

**Domain transfer:** data from one domain is often not reusable in another

**Structure:** Handcrafted structures are often necessary to make learning reliable

**Reproducibility problems** and recent improvements

- bAbI Dialogue tasks [Bordes et al., 2017] (ParlAI [Miller et al., 2018])
- PyDial (benchmarking) environment [Ultes et al., 2017]

# Revisiting Problems of RL-Based DM

**PyDial:** Open-source task-oriented statistical DM toolkit

- **Benchmarking environment:** Simulated environments in 3 domains and 6 different environments (different SER, user behaviour, masks...)
- New domains and environments easily implementable

A recent benchmarking study of different DM approaches  
[Casanueva et al., 2017]

- Domain complexity plays a crucial role, but...
- **Data-driven models still fall way short of hand-crafted policies...**
- Even in very constrained environments compared to real-world tasks

# Revisiting Problems of RL-Based DM

## Some Quantitative Evidence

		GP-Sarsa		DQN		A2C		eNAC		Handcrafted	
<i>Task</i>		Suc.	Rew.	Suc.	Rew.	Suc.	Rew.	Suc.	Rew.	Suc.	Rew.
Env. 1	CR	99.4%	<b>13.5</b>	93.9%	12.7	89.3%	11.6	94.8%	12.4	100.0%	14.0
	SFR	96.1%	11.4	65.0%	5.9	58.3%	4.0	94.0%	<b>11.7</b>	98.2%	12.4
	LAP	89.1%	9.4	70.1%	6.9	57.1%	3.5	91.4%	<b>10.5</b>	97.0%	11.7
Env. 2	CR	96.8%	<b>12.2</b>	91.9%	12.0	75.5%	7.0	83.6%	9.0	100.0%	14.0
	SFR	91.9%	<b>9.6</b>	84.3%	9.2	45.5%	-0.3	65.6%	3.7	98.2%	12.4
	LAP	82.3%	<b>7.3</b>	74.5%	6.6	26.8%	-5.0	55.1%	1.5	97.0%	11.7
Env. 3	CR	95.1%	11.0	93.4%	<b>11.9</b>	74.6%	7.3	90.8%	11.2	96.7%	11.0
	SFR	81.6%	6.9	60.9%	4.0	39.1%	-2.0	84.6%	<b>8.6</b>	90.9%	9.0
	LAP	68.3%	4.5	61.1%	4.3	37.0%	-1.9	76.6%	<b>6.7</b>	89.6%	8.7
Env. 4	CR	91.5%	9.9	90.0%	<b>10.7</b>	64.7%	3.7	85.3%	9.0	96.7%	11.0
	SFR	81.6%	7.2	77.8%	<b>7.7</b>	38.8%	-3.1	61.7%	2.0	90.9%	9.0
	LAP	72.7%	5.3	68.7%	<b>5.5</b>	27.3%	-6.0	52.8%	-0.8	89.6%	8.7
Env. 5	CR	93.8%	9.8	90.7%	10.3	70.1%	5.0	91.6%	<b>10.5</b>	95.9%	9.7
	SFR	74.7%	3.6	62.8%	2.9	20.2%	-5.9	74.4%	<b>4.5</b>	87.7%	6.4
	LAP	39.5%	-1.6	45.5%	0.0	28.9%	-4.7	75.8%	<b>4.1</b>	85.1%	5.5
Env. 6	CR	89.6%	8.8	87.8%	<b>10.0</b>	62.3%	3.5	79.6%	8.0	89.6%	9.3
	SFR	64.2%	2.7	47.2%	0.4	27.5%	-5.1	66.7%	<b>3.9</b>	79.0%	6.0
	LAP	44.9%	-0.2	46.1%	1.0	32.1%	-3.8	64.6%	<b>3.6</b>	76.1%	5.3
Mean	CR	94.4%	10.9	91.3%	<b>11.3</b>	72.8%	6.4	87.6%	10.0	96.5%	11.5
	SFR	81.7%	<b>6.9</b>	66.3%	5.0	38.2%	-2.1	74.5%	5.7	90.8%	9.2
	LAP	66.1%	4.1	61.0%	4.1	34.9%	-3.0	69.4%	<b>4.3</b>	89.1%	8.6
	ALL	80.7%	<b>7.3</b>	72.9%	6.8	48.6%	0.4	77.2%	6.7	92.1%	9.8

Table 4: Reward and success rates after 4000 training dialogues for the five policy models considered in this benchmark. Each row represents one of the 18 different tasks. The highest reward obtained by a data driven model in each row is highlighted.

# Dialogue Management

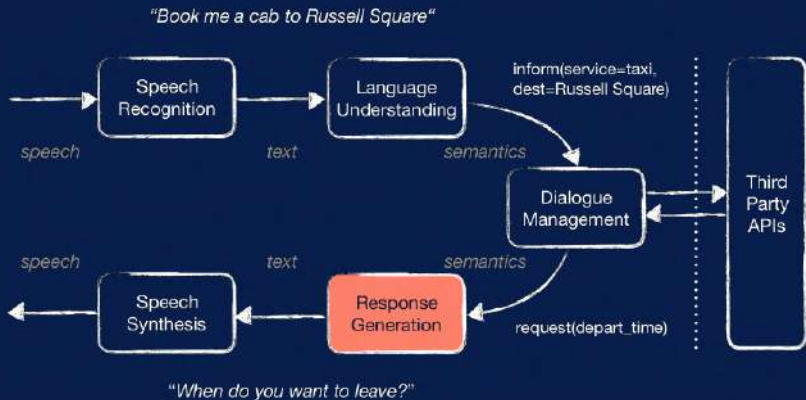
## Based on RL

- Machine Learning finds the solution from the data
- People care much about “learning”, but less so about the domain problems

Deep Reinforcement Learning:

	State	Action	Reward
AlphaGO	19x19 grid	Next move	win/loss
Dialogue	User query	Language response	satisfaction, success

# Natural Language Generation



# Natural Language Generation

## Problem Definition

“NLG is the process of deliberately constructing a natural language text in order to meet specified communicate goals.”

[McDonald, 1992]

**Input:** a communicative goal. E.g. non-linguistic representation of information

**Output:** natural language utterance (combined with graphics, tables, etc)

**Knowledge Required:** intended communication scope, domain knowledge, the actual language



# Natural Language Generation

## Problem Definition

- Meaning representation → natural language utterances

*Dialogue Act*

*Realisations*

*Inform(restaurant=Seven\_days, food=Chinese)*

*Seven days is a restaurant serving Chinese.*

*Seven days is a Chinese restaurant.*

- What do we care about? - naturalness and variation
- Other applications - summarisation, short-text report

# NLG Evaluation

Subjective: human judgement [Stent et al., 2005]

- **Adequacy:** correct meaning
- **Fluency:** linguistic fluency/naturalness
- **Readability:** fluency in the dialogue context
- **Variation:** multiple realisations of the same meaning

Objective: automatic evaluation measures

- Word overlap: BLEU, METEOR, ROUGE
- Embedding-based: greedy matching, embedding average
- Task-oriented: item error rate

**Big gap and anti-correlation between human judgements and automatic measures**

# NLG Evaluation

The gap between human perception and automatic metrics

[Stent, 2005; Liu et al., 2017; Sharma et al., 2017]

Correlation	Adequacy	Fluency
BLEU	0.388	-0.492

Real user trial is still the best way to evaluate NLG

...but it is inefficient and suffer from inter/intra-rating mismatch

# Template-Based NLG

- Define a set of rules to map meaning representation (i.e., frames) to natural language

```
confirm()           "Please tell me more about the product your are looking for."  
confirm(area=$V)    "Do you want somewhere in the $V?"  
confirm(food=$V)    "Do you want a $V restaurant?"  
confirm(food=$V,area=$W) "Do you want a $V restaurant in the $W."
```

...

- **Pros:** simple, error-free, easy to control
- **Cons:** time-consuming, poor scalability, repetitive

# Pipeline approach to NLG

## What to say?

- Content planning: pick contents and order them
- Dialogue policy handles this in SDS

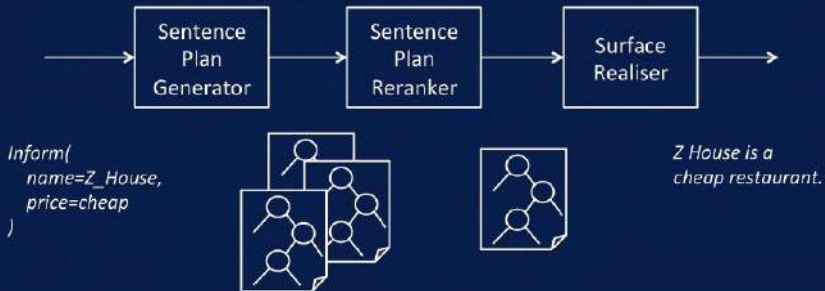
## How to say it?

- Sentence planning: get sentence (tree), put content in right places, lexicalisation, etc
- Surface realisation: creates linear text and ensures its grammar

# The "how to say it" Pipeline in NLG

[Walker et al., 2002]

## ■ Sentence planning + Surface realisation



# Statistical approaches to pipelined NLG

- Statistical sentence plan generator  
[Stent et al., 2009]
  - Statistical surface realizer  
[Dethlefs et al., 2013; Cuayahuitl et al., 2014]
  - Learning from unaligned data  
[Dušek and Jurčiček, 2015]
- 
- **Pros:** can model complex linguistic structures
  - **Cons:** heavily engineered, require extensive domain knowledge

# Sequential approaches to NLG

Class-Based NLG [Oh and Rudnicky, 2000]

## Classes

inform\_area, inform\_address, ..., request\_area, request\_address, ...

## Class-based Language Modeling

- $P(X|c) = \sum_t \log p(x_t|x_0, x_1, \dots, x_{t-1}, c)$
- $X^* = \arg \max P(X|c)$
- **Pros:** easy to implement/understand, simple rules
- **Cons:** computationally inefficient

## Further work

- ▶ Phrase-based NLG [Mairesse et al., 2010]
- ▶ ...

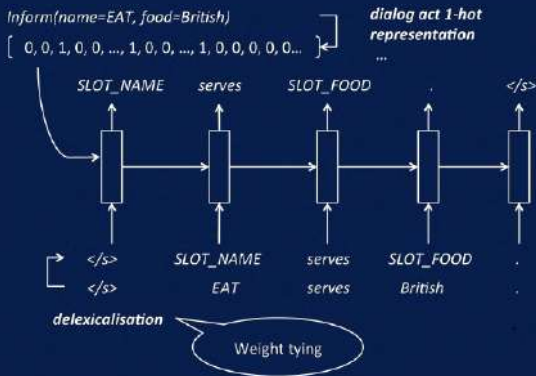


# RNN-Based LM NLG

[Wen et al., 2015]

“Deep” revolution has also transformed NLG

- **Distributed representations:** Generalisation
- **Recurrent connections:** Long-term dependencies
- **Conditional RNNs:** Flexibility and creativity
- **Attention mechanism:** Focus on relevant and meaningful parts



# RNN-Based LM NLG

## Handling Semantic Repetition/Missing

Empirically, **semantic repetition** is observed

- “EAT is a great **British** restaurant that serves **British food**.”
- “EAT is a **pet-friendly** place in the **cheap** price range. It is **affordable** and also **allows pets**.”

How to mitigate this problem?

- Post-processing rules [Oh and Rudnicky, 2000]
- Gating mechanism [Wen et al., 2015]
- Attention [Mei et al., 2016; Wen et al., 2016; Tran and Nguyen, 2017]

# RNN-Based NLG

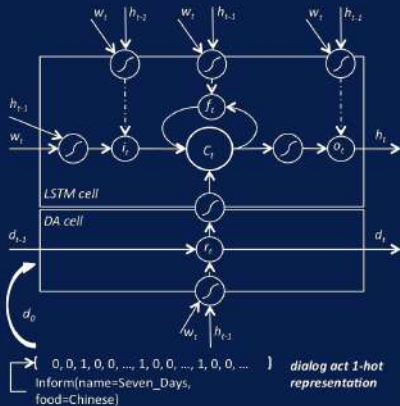
## Learning to Control Gates

Learn to open/close the LSTM gates based on **generation history** [Wen et al., 2015]

### Cost function

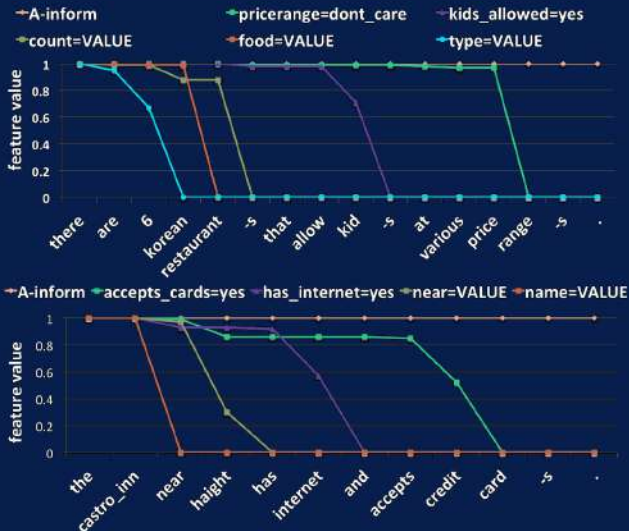
$$\mathcal{L}(\theta) = - \sum_t \mathbf{y}_t^T \log \mathbf{p}_t \\ + \|\mathbf{d}_T\| \\ + \sum_{t=0}^{T-1} \eta \xi \|\mathbf{d}_{t+1} - \mathbf{d}_t\|$$

- 1<sup>st</sup> term : Log-likelihood
- 2<sup>nd</sup> term: make sure rendering all the information needed
- 3<sup>rd</sup> term: close only one gate at each time step.



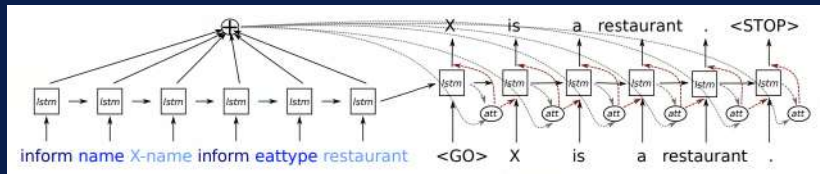
# RNN-Based NLG

## Feature Visualisation



# (Attentive) Seq2Seq for NLG

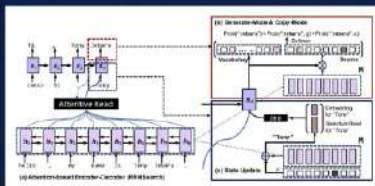
Encode Dialogue Act, Seq2Seq for generation [ [Dušek and Jurčiček, 2016](#), [Wen et al., 2015](#); ...]



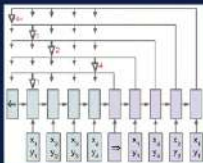
# Recent Work in Neural NLG

Re-using info from the input directly: CopyNet and PointerNet [Gu et al., 2016; Gulcehre et al., 2016; Cao et al., 2016; He et al., 2016; Miao et al., 2016; Ling et al., 2017 ...]

Copy Network

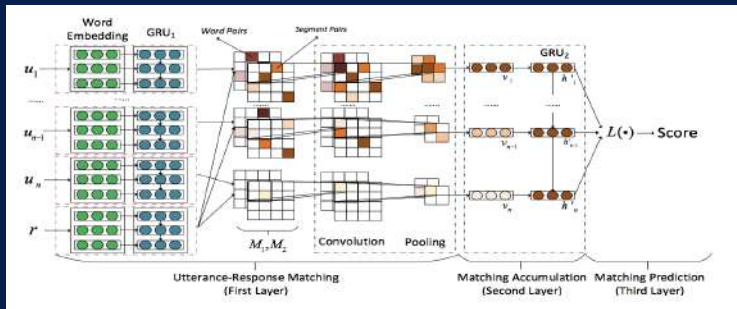


Pointer Network



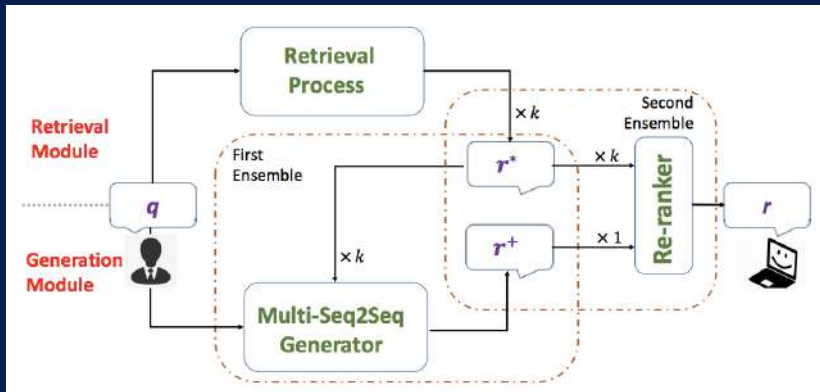
# Recent Work in Neural NLG

Retrieval-based NLG: select from a set of predefined responses  
[Lowe et al., 2015; Yan et al., 2016; **Wu et al., 2017**]



# Recent Work in Neural NLG

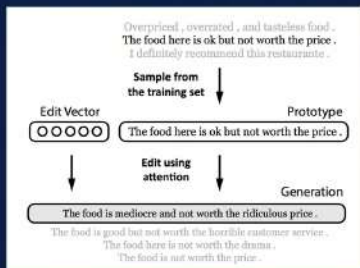
Combining retrieval-based and generative NLG [Song et al., 2016]





# Recent Work in Neural NLG

## Generating Sentences by Editing Prototypes [Gua et al., 2017]



	Example 1	Example 2
Context	I've had better service at <org>.	my daughter actually looks forward to going to the dentist!
Edit	+ worst - worse	+ his - her
Result	= best service i've had at <org>.	= my son actually looks forward to going to the dentist!

# Recent Work in Neural NLG

- Controlled text generation based on VAE/GANs [Wen et al., 2017; Li et al., 2017; Hu et al., 2017, Subramanian et al., 2017; Fedus et al., 2018]
- Slot-value informed Seq2Seq models [Sharma et al., 2017; Nayak et al., 2017]

Check also a recent survey paper on NLG focused on more traditional approaches [Gatt and Krahmer, 2018]

# Learning Objective for NLG

Closely related to the NLG evaluation

- Maximising the log-likelihood only optimises the perplexity.

$$\ell(\theta) = - \sum_{t=1}^T \log p(y_t | X, y_{<t}; \theta)$$

Other approaches

- Discriminative Training [Wen et al., 2016]
  - ▶ Separating correct examples from competing incorrect ones
- Reinforcement Learning [Li et al., 2016]
  - ▶ Additional info as reward: Mutual Information, Semantic Coherence, Information Flow...

# Corpus and Data for NLG

Statistical models are reflecting what's in the data

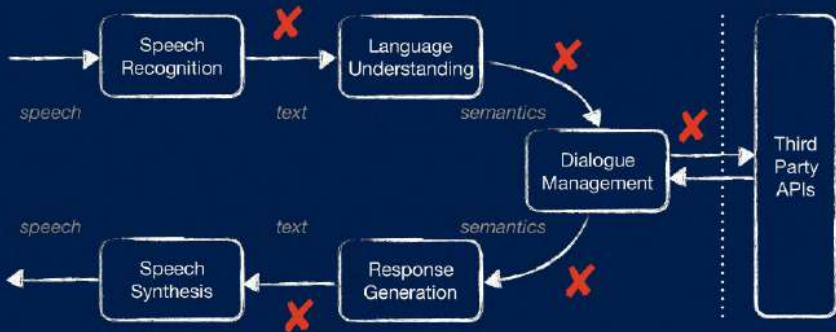
- Data quality decides the quality of the NLG system
- Do not learn from templates or synthetic data - you are just reverse engineering the design
- Collect your own real user data! [Wen et al., 2016; Shah et al., 2018; ...]

# Short Summary on NLG

- Evaluating NLG is hard. The best way is *human evaluation*
- In product, template-based NLG is still most common
- Learning-based NLG (+ linguistic knowledge) are promising, but be careful of the objective and data
- NN-based NLG is a conditional neural LM that learns *realisation* and *semantic alignments* jointly
- Recent trends: adversarial modeling, E2E learning

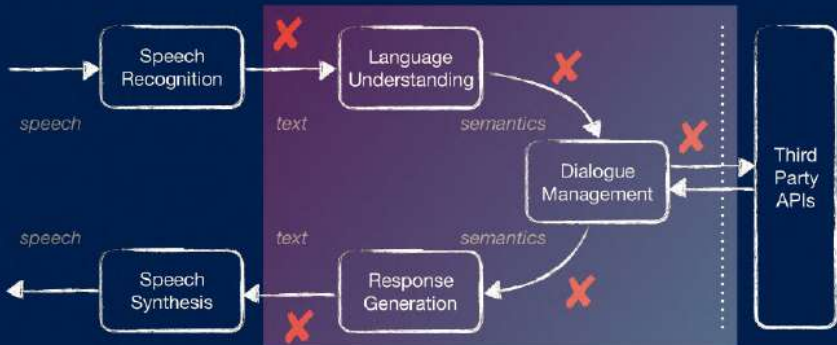
# End to End Dialogue Modelling

Why end to end?



# End to End Dialogue Modelling

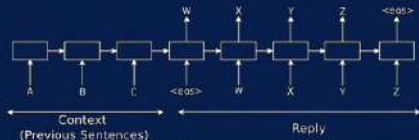
Why end to end?



# Sequence-to-Sequence Conversational Model

## Social Chatbots (Chit-chat dialogue systems)

[Vinyals et Le, 2015]



Human: what do you think about terla ?  
Machine: he 's a good conductor .  
Human: what do you think about bill gates ?  
Machine: he 's a good man .  
Human: What do you think about messi ?  
Machine: he 's a great player .  
Human: what do you think about cleopatra ?  
Machine: oh , she 's very regal .  
Human: what do you think about england during the reign of elizabeth ?  
Machine: it was a great place .

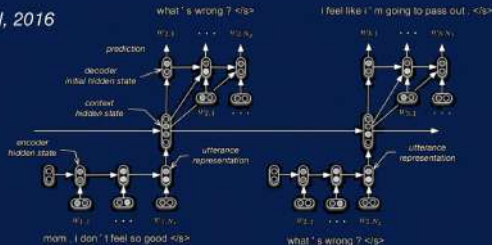
Human: what is the purpose of life ?  
Machine: to serve the greater good .  
Human: what is the purpose of living ?  
Machine: to live forever .  
Human: what is the purpose of existence ?  
Machine: to find out what happens when we get to the planet earth .  
Human: where are you now ?  
Machine: i 'm in the middle of nowhere .  
Human: what is the purpose of doing ?  
Machine: to have a life .  
Human: what is the purpose of being intelligent ?  
Machine: to find out what it is .  
Human: what is the purpose of emotions ?  
Machine: i don 't know .

- Ignore the structure of the problem.
- A simple sequence-to-sequence model to learn a context-to-response mapping.
- Maximising a conditional probability  $p(\text{response}|\text{context})$
- Large model with huge data.
- Can learn to generate interesting responses.



# Hierarchical RNN for Dialogue

Serban et al, 2016



Reference ( $U_1, U_2$ )	MAP	Target ( $U_3$ )
$U_1$ : yeah , okay .	i ' ll see you tomorrow .	yeah .
$U_2$ : well , i guess i ' ll be going now ,		
$U_1$ : oh . <continued utterance> oh .	i don ' t know .	oh .
$U_2$ : what ' s the matter , honey ?		
$U_1$ : it ' s the cheapest .	no , it ' s not .	they ' re all good , sir .
$U_2$ : then it ' s the worst kind ?		
$U_1$ : <person> ! what are you doing ?	what are you doing here ?	what are you that crazy ?
$U_2$ : shut up ! c ' mon .		

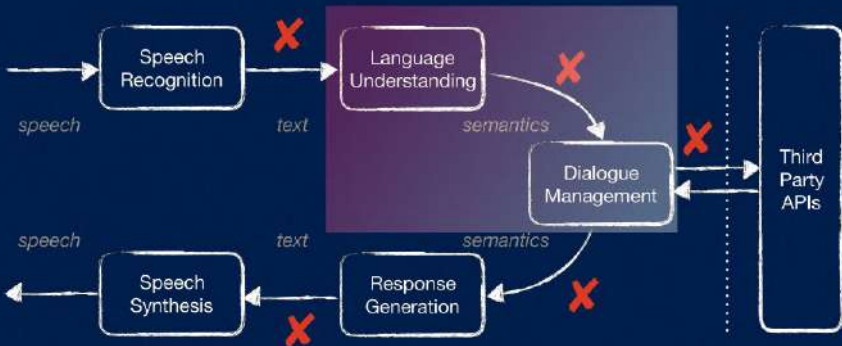
**Problems:** Generic, difficulty keeping coherence, lack of integration into KBs or 3rd party services... **not very useful for business use cases** (task oriented dialogues)

# Task Oriented End-to-End Dialogue Modelling

## Joint NLU and Policy Learning

No need to define / annotate belief states

[Zhao and Eskenazi, 2016]

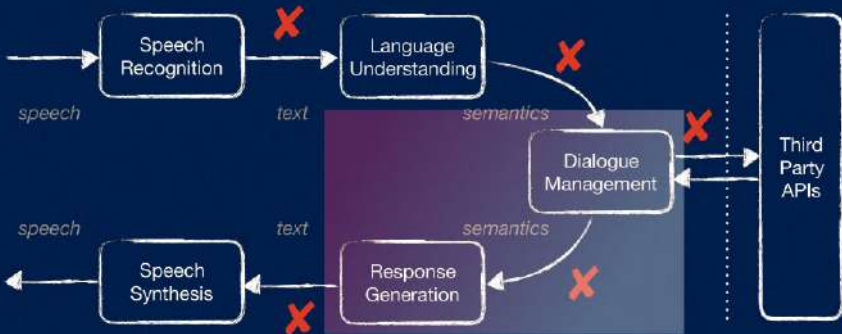


# Task Oriented End-to-End Dialogue Modelling

## Joint Policy and NLG Learning

No need to define / annotate dialogue acts

[Wen et al, 2017]

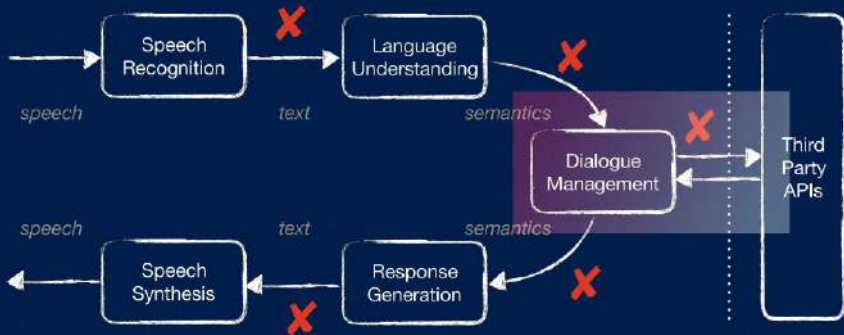


# Task Oriented End-to-End Dialogue Modelling

## Joint Policy and DB query learning

No need to define DB query heuristics

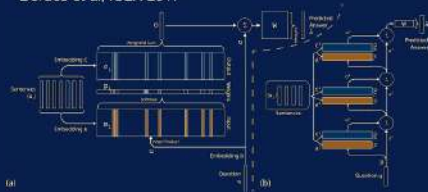
[Dhingra et al., 2017]



# Task Oriented End-to-End Dialogue Modelling

## Memory Networks for Dialogue Modeling

Bordes et al, ICLR 2017



- Without using the structure of dialogue systems.
- Apply attention mechanism to search for relevant information in the context.
- Aggregate the current user query with the retrieved context to make predictions.
- Predictions include API calls and response selection from a candidate pool.
- Task success rate is only around 41.1% when trained on 2000 dialogues.

Time	Location	Dialog History	Resp #1	Resp #2
1	User	hey can you	.180	.095
2	User	could you check if i can get a reservation at <corp> <date> for brunch	.339	.178
3	User	<numbers> people	.197	.142
4	User	<address>	.187	.167
5	Bot	hi <name> sorry unfortunately <corp> is fully booked for <date> and there's <numbers> people on the waiting list	.325	.107
User input		when's the earliest availability		
Correct answer		i'll check		
Pred. answer #1		i'm in it		[Incorrect]
Pred. answer #2		i'll find out		[Incorrect]
Pred. answer #3		i'll take a look		[Incorrect]
Pred. answer #4		i'll check		[Correct]
Pred. answer #5		i'll check into it		[Incorrect]

more related work [Liu and Perez, 2016; Eric and Manning, 2017; Seo et al., 2016]

# Where are we in End-to-End Dialogue Modelling?

Most E2E approaches are presented in very simple domains and non-reproducible environments: **the problems adapted to the model** and not vice versa

bAbI dialogue tasks (and parlAI) & DSTC6 will make E2E dialogue research more reproducible

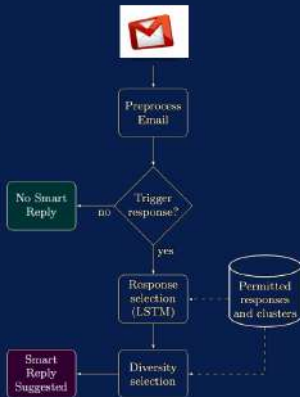
Generative E2E dialogue modelling is extremely hard to get it work, and still seems far from real world applications. But what about discriminative models?

# A Working-Example: Google Smart Reply

- Response selected from a whitelist.
- LSTM is used for sentence scoring:

$$P(r_1, \dots, r_m | o_1, \dots, o_n) = \prod_{i=1}^m P(r_i | o_1, \dots, o_n, r_1, \dots, r_{i-1})$$

- Following the Seq2Seq paradigm.
- Inference is done by:
  - Organizing responses as a trie
  - Beam search



Kannan et al, 2016

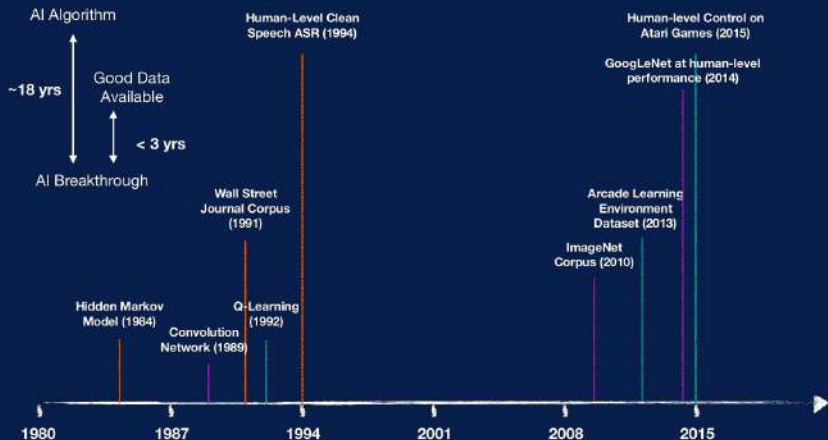
**Data! Data! Data!**



# The Importance of Data

Data is the core of ML algorithms

<http://www.spacemachine.net/views/2016/3/datasets-over-algorithms>



# Getting Data for Training Dialogue Systems

Conversational data is hard to get

Reddit has a large volume of conversational data...

- ...Well, for chit-chat dialogue development



But how do we get **high-quality domain-specific** data sets?

- A pizza delivery agent?
- A taxi booking service?
- A flight booking agent?

# Wizard-of-Oz (WOZ) Data Collection

Online parallel version of WOZ on MTurk (or any other platform)

- Randomly hire a worker to be user/wizard
- Task: Enter an appropriate response for one turn
- Repeat the process until all dialogues are finished

Task 02004: You are looking for and it should serve **gastropub food**. You don't care about the **price range**. You want to know the **address**.

Info Desk : Hello, welcome to the Cambridge restaurant system. You can ask for restaurants by area, price range or food type. How may I help you?  
Customer : I want a gastropub food.  
Info Desk : There are 4 restaurants serving gastropub food, what price range do you want?

Instructions

Customer : (Your response)

I don't care about the price range, just give me the address please.

Submit the HIT

# Wizard-of-Oz (WOZ) Data Collection

## Example wizard page

info Desk: Hello, welcome to the Cambridge restaurant system. You are asking for restaurants by area, price range or food type. How may I help you?

Customer: I want a gastropub food.

info Desk: There are 6 restaurants serving gastropub food, what price range do you want?

Customer: I don't care.

**Next step**

Please identify the following answers based on the latest customer response:

- What does user want?  
What is the food type the user wants? gastropub  
What is the area the user wants? not mentioned  
What is the price range the user wants? don't care
- What does user ask?  
Is the user asking for food type of an offered result? No ☐  
Is the user asking for price range of an offered result? No ☐  
Is the user asking for area of an offered result? No ☐  
Is the user asking for a parade of an offered result? No ☐  
Is the user asking for phone number of an offered result? No ☐  
Is the user asking for address of an offered result? No ☐  
Is the user mentioning any restaurant names? No ☐

**Submit**

info Desk: (User responses)

would recommend top rated (best), a great gastropub restaurant in the centre, do you want their phone number? and of course? **Submit the list**

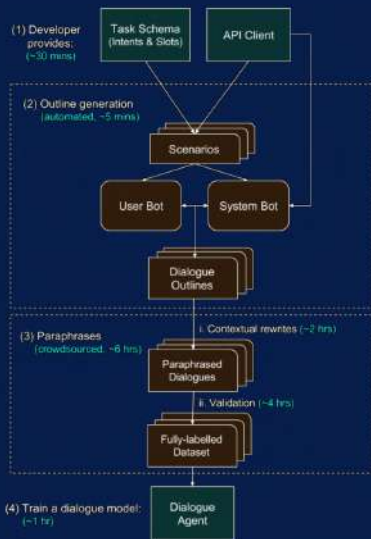
Name	Food	Area	Price Range	Phone	Address	Postcode
backroom bistro	gastropub	centre	expensive	01223 300009	2 Burton Street City Centre	GB 1 1 GA
cycl standard	gastropub	east	expensive	01223 241877	205 New Road City Centre	GB 1 3 NL
the new jazz kitchen and bar	gastropub	centre	moderate	01223 368871	Corner Exchange Street	GB 2 1 QF
travelling and reduce	gastropub	centre	expensive	—	94 - 96 Green Street	GB 2 1 JU
—	gastropub	—	—	—	—	—

Showing 1 to 4 of 4 entries (Sorted from 1 to last entries)

Previous **1** Next

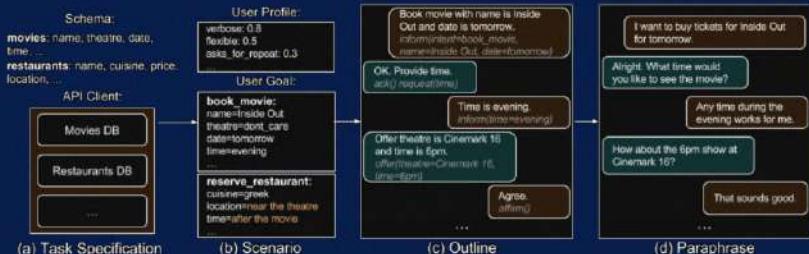
# Getting Data for Training Dialogue Systems

Machines Talking to Machines (M2M) [Shah et al.; NAACL 2018]



# Getting Data for Training Dialogue Systems

Machines Talking to Machines (M2M) [Shah et al.; NAACL 2018]



# Getting Data for Training Dialogue Systems

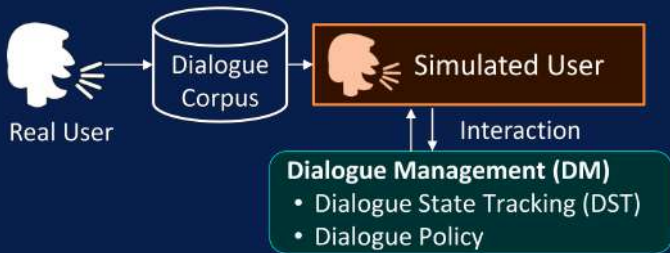
M2M: Reduced annotation complexity

Crowdsourcing simply as a **paraphrasing task**

Original conversation	Paraphrased conversation
<b>Assistant:</b> greetings	Hello
<b>User:</b> greetings AND buy movie tickets AND date is Wednesday	I would like to buy movie tickets for Wednesday.
<b>Assistant:</b> provide preference for: theatre name and time and number of tickets and movie	What theatre, movie, time, and number of tickets would you like?
<b>User:</b> unable to understand, must re-phrase your message	Sorry, I didn't understand. Could you repeat that?
<b>Assistant:</b> provide preference for: theatre name and time and number of tickets and movie	How many tickets would you like and what is the theatre, movie, and showtime?
<b>User:</b> movie is The Magnificent Seven and time is 4:00 pm and theatre name is Century 20 Great Mall and number of tickets is 6	I need 6 tickets for the 4:00 pm showing of The Magnificent Seven at the Century 20 Great Mall
<b>Assistant:</b> confirm the following: time is 4:00 pm	That was for 4:00 pm, right?
<b>User:</b> accept time is 4:00 pm	Yes, 4:00 pm.
<b>Assistant:</b> purchase confirmed with given details movie is The Magnificent Seven and number of tickets is 6 and time is 4:00 pm and date is March 06 and theatre name is Century 20 Great Mall	I have purchased 6 tickets for the Magnificent Seven at 4:00 o'clock on March 8 at the Century 20 Great Mall.
<b>User:</b> thank you AND goodbye	Thanks, bye.

# Data Simulation

**Goal:** generate natural and reasonable conversations to enable reinforcement learning for exploring the policy space



## Approaches:

- Rule-based (hand-crafted by experts) [Li et al., 2016]
- Learning-based [Schatzmann et al., 2006; El Asri et al., 2016; Shah et al., 2018]



# Data Simulation

## Specifying the dialogue flow before collecting the actual data

- How to cover all possible use cases?
- How to maximise coverage of different flows and slot-value pairs?
- How to quickly build simulators for new domains?

```
Turn 0.
=====
current_user_template: "I want price range is cheap\n AND price range is NOT moderate."
last_system_template: "Greetings!"
user_concept_highlights: [(price range="cheap") [-1:-1], (price range!="moderate") [-1:-1]]
system_concept_highlights:[]
dialogue_turn {
  dialogue_id: "2e091838-6d2a-4ed5-bf06-caac5cde4d0e"
  last_system_act: GREETING()
  user_action: INFORM
  slu_labels: (price range="cheap", price range!="moderate")
  state_labels: (price range="cheap", price range!="moderate")
}

Turn 1.
=====
current_user_template: "I want has_wheelchair_access is True\n AND cuisine is Hot Dogs."
last_system_template: "How about name is Boyce Da Roca\n AND price range is cheap?\n Anything else?"
user_concept_highlights: [(accessible_wheelchair=True) [-1:-1], (cuisine="Hot Dogs") [-1:-1]]
system_concept_highlights: [(name="Boyce Da Roca") [-1:-1], (price range="cheap") [-1:-1]]
dialogue_turn {
  dialogue_id: "2e091838-6d2a-4ed5-bf06-caac5cde4d0e"
  turn_index: 1
```

# WOZ vs M2M

## WOZ data collection:

- No explicit dialogue act annotations
- System policy can be learned directly from the data
- Interesting and diverse system behaviours (but long-tail...)
- Not easy to control the system behaviour
- Black box during system development
- User task simulation has to be created
- Complex UI for data annotation
- Difficult to scale to large domains (for now)

# WOZ vs M2M

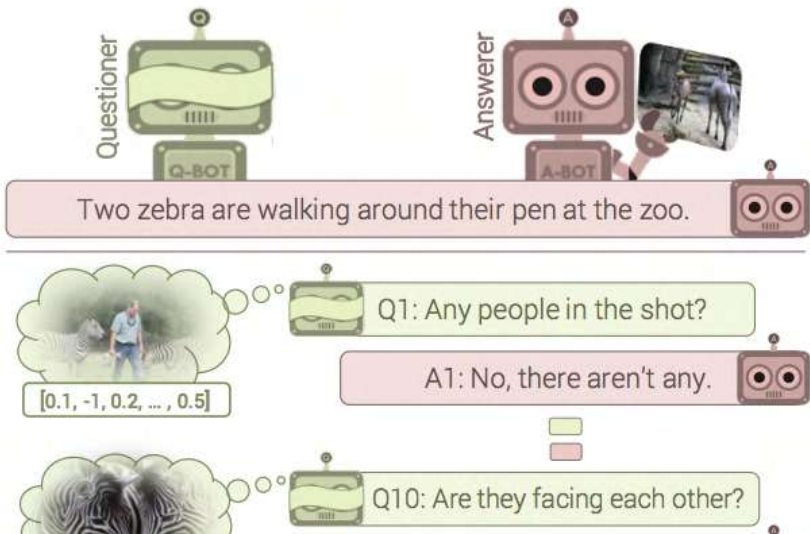
## M2M data collection:

- Full control over the dialogue flow
- Paraphrase data collection UI is simpler to build
- Easier to engineer particular behaviours
- Crowdsourcers don't have to label data
  
- Notation of dialogue acts while developing the dialogue flow
- Simulating both users and the hand-crafted system
- No interesting system behaviours learned; everything is expected

# Trends, Challenges, Future


# Cooperative Visual Dialog using DRL

[Das et al., 2016, 2017]




# Knowledge-Grounded Conversational AI


[Ghazvininejad et al., 2017]




"Consistently the best omakase in San Francisco." (2 Tips)




"Probably the best sushi in San Francisco." (2 Tips)



"... they were out of the kaisui sari by the time we ate, but the bafun uni is..." (2 Tips)




"Amazing sushi tasting from the chefs of Sushi Ran" (2 Tips)

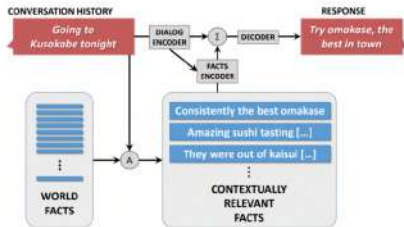


San Francisco

○ Kusakabe



**User input:** Going to Kusakabe tonight.  
**Neural model:** Have a great time!  
**Human:** You'll love it! Try omakase, the best in town.



A: Looking forward to trying @pizzalibretto tonight! my expectations are high.

B: Get the rocco salad. Can you eat calamari?

A: Anyone in Chi have a dentist office they recommend? I'm never going back to [...] and would love a reco!

B: Really looved Ora in Wicker Park.

A: I'm at California Academy of Sciences

B: Make sure you catch the show at the Planetarium. Tickets are usually limited.

A: I'm at New Wave Cafe.

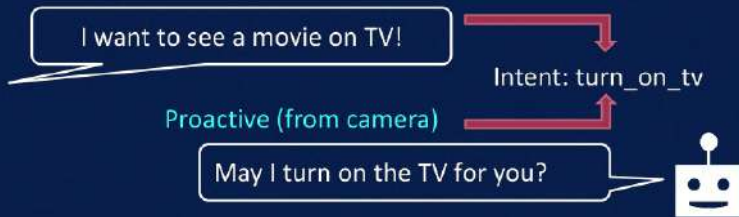
B: Try to get to Dmitri's for dinner. Their pan fried scallops and shrimp scampi are to die for.

A: I just bought: [...] 4.3-inch portable GPS navigator for my wife, shh, don't tell her.

B: I heard this brand loses battery power.

# Multi-Modal Conversational AI

## Intent Understanding from the Visual Context?



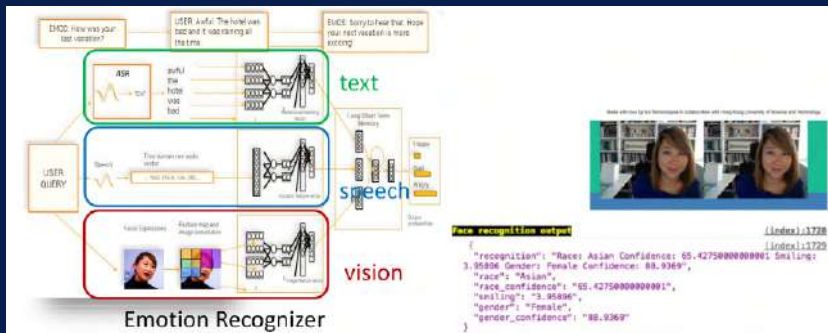
**Proactive behaviour:** understand user intents to initiate the dialogues

# Multi-Modal Conversational AI

## Emotion Recognition and Empathy?

Integrate an emotion recognition and empathy module

- Multi-modal emotion recognition
- Emotion-aware response generation

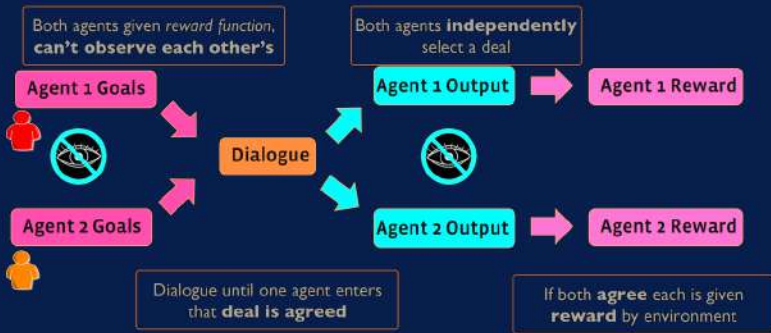




# (The Art of) Negotiation Dialogue

[Lewis et al., 2017; Keizer et al., 2017]

## Framework



- Both linguistic and reasoning problem
- Negotiations are hard and awkward, incentive to withhold information
- There is no simple solution, but easy to evaluate

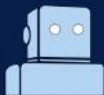
# Alexa Skills Challenge

Adding new abilities to Amazon Alexa



- An augmented reality virtual pet
- The dealer in poker
- Trip planner based on FourSquare, Google Places info
- Full voice control over Chrome browser
- Tweeting using voice
- Live ice hockey updates...

# Industry Landscape



Chatfuel



KIT.ai



Converse AI



Maluuba

A Microsoft company

AUTOMAT



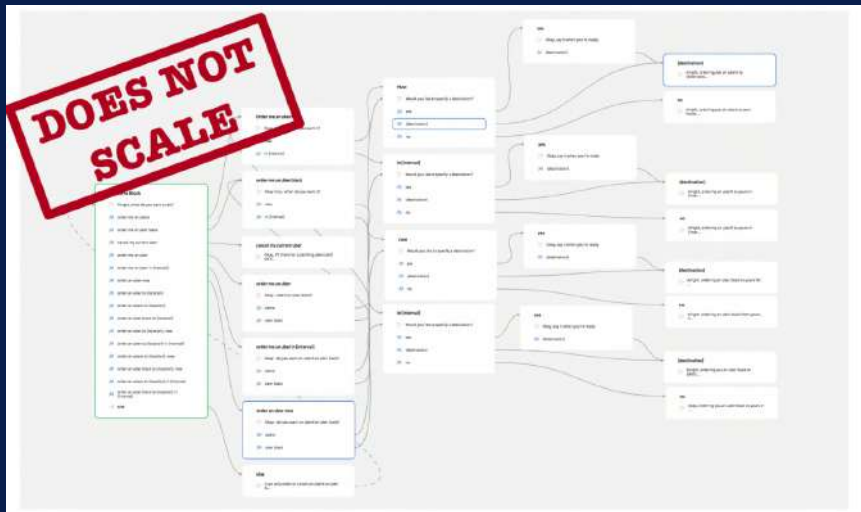
Dialogflow



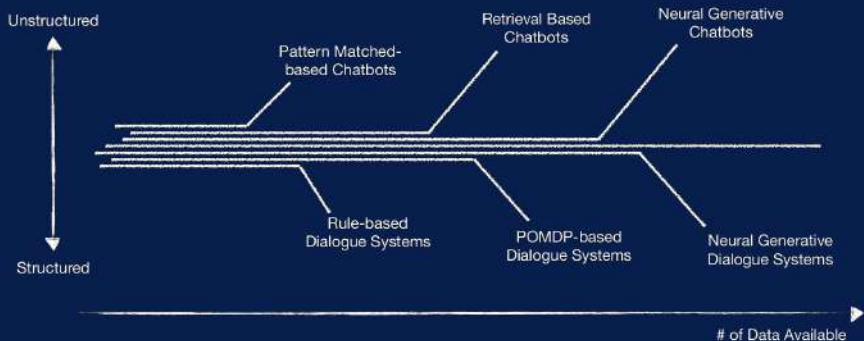
Ozlo



## Current Industry Practice: Flowcharts

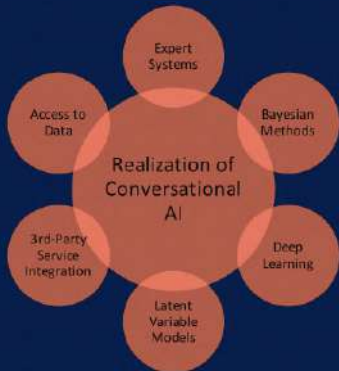


# A Brief Summary of Conversational AI Approaches



# Conclusion

## A collective effort!



Realizing Conversational AI requires efforts from multiple parties:

- 3rd-Party Service Providers.
- Conversational UX/UI Developers.
- Machine Learning Scientists/Engineers.

Realizing Conversational AI also requires a hybrid of approaches to tackle different parts of the pipeline:

- Expert programmed algorithms to handle domain-specific **business logics**.
- Different machine learning models to handle **prediction** and **uncertainty estimation**.
- Integration to **3rd-party services** to provide meaningful applications.
- Access to data to learn from human interactions.

# Conclusion

Conversational AI is a collective effort of **multiple parties** and requires a mix of **multiple methods**

Machine Learning is...

- a perfect tool for language understanding
- an OK method for response selection/generation
- currently not a way to go for dialogue decision-making and database integration (we need more and more high-quality data)

Being **practical but open-minded** in choosing approaches is the way to build robust and truly useful conversational AI of the future

# Existing Tools



## Complex Flowcharts

- Entirely hand-crafted turn-by-turn dialogue flow design
- Developers have to pre-specify all possible use cases



## Weak Language Understanding

- Keyword spotting or basic machine learning techniques
- Users have to know which words to use to invoke each command



## Learning From Data is Impossible

- Real-user data cannot be used to optimise user experience
- Handling complex dialogue context with flowcharts is near-impossible

- ✗ **Hard to Build**
- ✗ **Unreliable**
- ✗ **Not Scalable!**

POLYAI



The Platform for Conversational AI



## Dynamic Flow Design

- Plug-and-play policy design with composable pre-built blocks
- Machine learning + available overrides for app-specific behaviour



## Deep Learning for NLP

- state-of-the-art deep learning models for robust language understanding and lexically rich response generation



## High-Quality Data Collection

- On-demand data collection via our proprietary crowdsourcing platform
- Dialogue quality control is based on our expertise in crowdsourcing

- ✓ **Easy to Build**
- ✓ **Robust**
- ✓ **Scalable!**



# A Not-So-Far-Away Future

*"Hey PolyBot, could you order a pizza home for me. I will be back around 7.30. You know my preference, right?"*



*"Hey PolyBot, what is the status of my order? Could you tell PizzaHut that I'm going to be 30 minutes late?"*



*"[Notification] Your Margarita order has arrived. Please give a 1-5 star feedback on our delivery service."*



# Useful Software I

**PyDial:** open-source toolkit with implementations of statistical approaches for all dialogue system modules

<http://dialogue.mi.eng.cam.ac.uk/index.php/pydial/>

**ParlAI:** “A unified platform for sharing, training, and evaluating dialog models across many tasks”

<http://parl.ai/>

**DeepPavlov:** another open-source conversational AI library

<https://github.com/deepmipt/DeepPavlov>

**NBT:** Implementation of the fully statistical NBT DST model

<https://github.com/nmrksic/neural-belief-tracker>

**Different DST Models** available on GitHub

<https://github.com/CallumMain/DNN-DST>

<https://github.com/voicy-ai/DialogStateTracking>

# Useful Software II

**NNDial:** end-to-end trainable task-based systems

<https://github.com/shawnwun/NNDIAL>

**RNNLG:** NLG using RNNs

<https://github.com/shawnwun/RNNLG>

**Neural Dialogue Generation:** implementations of NLG methods from the Stanford NLP Group

<https://github.com/jiweil/Neural-Dialogue-Generation>

**Other open-source NLG software**

<https://github.com/simplenlg/simplenlg>

<http://www.fb10.uni-bremen.de/anglistik/langpro/kpml/README.html>

<http://nlp.cs.aueb.gr/software.html>

# Useful Software III

**Deep RL:** a simple deep RL dialogue system

<https://github.com/cuayahuitl/SimpleDS>

**Attract-Repel:** tool for word vector space specialisation

<https://github.com/nmrksic/attract-repel>

**DSTC:** data and scripts from the latest campaigns

<https://github.com/oplatek/sds-tracker>

<https://github.com/hkhpublish/dstc4>

<https://github.com/seokhwankim/dstc5>

<https://github.com/perezjln/dstc6-goal-oriented-end-to-end>

<https://github.com/dialogtekgeek/DSTC6-End-to-End-Conversation-Modeling>

# (A Selection of) References I



Layla El Asri, Jing He, and Kaheer Suleman. 2016.  
A sequence-to-sequence model for user simulation in spoken dialogue systems.  
*CoRR*, [abs/1607.00070](#).



Layla El Asri, Romain Laroche, and Olivier Pietquin. 2013.  
Reward shaping for statistical optimisation of dialogue management.  
In *Proceedings of SLSP*, pages 93–101.



Antoine Bordes and Jason Weston. 2017.  
Learning end-to-end goal-oriented dialog.  
In *Proceedings of ICLR*.



Paweł Budzianowski, Stefan Ultes, Pei-Hao Su, Nikola Mrkšić, Tsung-Hsien Wen, Iñigo Casanueva, Lina M. Rojas Barahona, and Milica Gasic. 2017.  
Sub-domain modelling for dialogue management with hierarchical reinforcement learning.  
In *Proceedings of SIGDIAL*, pages 86–92.

# (A Selection of) References II



Iñigo Casanueva, Pawel Budzianowski, Pei-Hao Su, Nikola Mrkšić, Tsung-Hsien Wen, Stefan Ultes, Lina Maria Rojas-Barahona, Steve J. Young, and Milica Gašić. 2017.

A benchmarking environment for reinforcement learning based task oriented dialogue management.

*CoRR*, [abs/1711.11023](#).



Iñigo Casanueva, Pawel Budzianowski, Pei-Hao Su, Stefan Ultes, Lina Maria Rojas-Barahona, Bo-Hsiang Tseng, and Milica Gasic. 2018.

Feudal reinforcement learning for dialogue management in large domains.



Nina Dethlefs, Helen Hastie, Heriberto Cuayáhuitl, and Oliver Lemon. 2013.

Conditional random fields for responsive surface realisation using global features.

In *Proceedings of ACL*, pages 1254–1263.



Bhuwan Dhingra, Lihong Li, Xiujun Li, Jianfeng Gao, Yun-Nung Chen, Faisal Ahmed, and Li Deng. 2017.

Towards end-to-end reinforcement learning of dialogue agents for information access.

In *Proceedings of ACL*, pages 484–495.

# (A Selection of) References III



Ondrej Dusek and Filip Jurcícek. 2016a.  
A context-aware natural language generator for dialogue systems.  
In *Proceedings of SIGDIAL*, pages 185–190.



Ondrej Dusek and Filip Jurcícek. 2016b.  
Sequence-to-sequence generation for spoken dialogue via deep syntax trees and strings.  
In *Proceedings of ACL*.



Ondřej Dušek and Filip Jurcicek. 2015.  
Training a natural language generator from unaligned data.  
In *Proceedings of ACL*, pages 451–461.



Mehdi Fatemi, Layla El Asri, Hannes Schulz, Jing He, and Kaheer Suleman. 2016.  
Policy networks with two-stage training for dialogue systems.  
In *Proceedings of ACL*, pages 101–110.



Pascale Fung, Dario Bertero, Yan Wan, Anik Dey, Ricky Ho Yin Chan, Farhad Bin Siddique, Yang Yang, Chien-Sheng Wu, and Ruixi Lin. 2016.  
Towards empathetic human-robot interactions.  
In *Proceedings of CICLING*, pages 173–193.

# (A Selection of) References IV



Milica Gasic, Catherine Breslin, Matthew Henderson, Dongho Kim, Martin Szummer, Blaise Thomson, Pirros Tsiakoulis, and Steve Young. 2013. POMDP-based dialogue manager adaptation to extended domains. In *Proceedings of SIGDIAL*, pages 214–222.



Milica Gašić, Nikola Mrkšić, Lina Maria Rojas-Barahona, Pei-Hao Su, Stefan Ultes, David Vandyke, Tsung-Hsien Wen, and Steve J. Young. 2017. Dialogue manager domain adaptation using Gaussian process reinforcement learning. *Computer Speech & Language*, 45:552–569.



Milica Gašić, Nikola Mrkšić, Pei-hao Su, David Vandyke, Tsung-Hsien Wen, and Steve J. Young. 2015. Policy committee for adaptation in multi-domain spoken dialogue systems. In *Proceedings of ASRU*, pages 806–812.



Hado van Hasselt, Arthur Guez, and David Silver. 2016. Deep reinforcement learning with double Q-learning. In *Proceedings of AAAI*, pages 2094–2100.



# (A Selection of) References V



Matthew Henderson, Rami Al-Rfou, Brian Strope, Yun-Hsuan Sung, László Lukács, Ruiqi Guo, Sanjiv Kumar, Balint Miklos, and Ray Kurzweil. 2017. Efficient natural language response suggestion for smart reply. *CoRR*, abs/1705.00652.



Matthew Henderson, Blaise Thomson, and Steve J. Young. 2013. Deep neural network approach for the Dialog State Tracking Challenge. In *Proceedings of SIGDIAL*, pages 467–471.



Matthew Henderson, Blaise Thomson, and Steve J. Young. 2014a. Robust dialog state tracking using delexicalised recurrent neural networks and unsupervised adaptation. In *Proceedings of SLT*, pages 360–365.



Matthew Henderson, Blaise Thomson, and Steve J. Young. 2014b. Word-based dialog state tracking with recurrent neural networks. In *Proceedings of SIGDIAL*, pages 292–299.



Jiwei Li, Will Monroe, Alan Ritter, Dan Jurafsky, Michel Galley, and Jianfeng Gao. 2016. Deep reinforcement learning for dialogue generation. In *Proceedings of EMNLP*, pages 1192–1202.

# (A Selection of) References VI



Jiwei Li, Will Monroe, Tianlin Shi, Sébastien Jean, Alan Ritter, and Dan Jurafsky. 2017.

Adversarial learning for neural dialogue generation.

In *Proceedings of EMNLP*, pages 2157–2169.



Chia-Wei Liu, Ryan Lowe, Iulian Serban, Mike Noseworthy, Laurent Charlin, and Joelle Pineau. 2016.

How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation.

In *Proceedings of EMNLP*, pages 2122–2132.



Ryan Lowe, Michael Noseworthy, Iulian Vlad Serban, Nicolas Angelard-Gontier, Yoshua Bengio, and Joelle Pineau. 2017.

Towards an automatic Turing test: Learning to evaluate dialogue responses.

In *Proceedings of ACL*, pages 1116–1126.



Andrea Madotto, Chien-Sheng Wu, and Pascale Fung. 2018.

Mem2Seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems.

*CoRR*, abs/1804.08217.

# (A Selection of) References VII



François Mairesse, Milica Gasic, Filip Jurčicek, Simon Keizer, Blaise Thomson, Kai Yu, and Steve J. Young. 2010.

Phrase-based statistical language generation using graphical models and active learning.

In *Proceedings of ACL*, pages 1552–1561.



Hongyuan Mei, Mohit Bansal, and Matthew R. Walter. 2016.

What to talk about and how? Selective generation using LSTMs with coarse-to-fine alignment.

In *Proceedings of NAACL-HLT*, pages 720–730.



Grégoire Mesnil, Yann Dauphin, Kaisheng Yao, Yoshua Bengio, Li Deng, Dilek Hakkani-Tur, Xiaodong He, Larry Heck, Dong Yu, and Geoffrey Zweig. 2015.

Using recurrent neural networks for slot filling in spoken language understanding. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(3):530–539.



Alexander Miller, Will Feng, Dhruv Batra, Antoine Bordes, Adam Fisch, Jiasen Lu, Devi Parikh, and Jason Weston. 2017.

ParlAI: A dialog research software platform.

In *Proceedings of EMNLP: Demos*, pages 79–84.

## (A Selection of) References VIII



Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin A. Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. 2015.

Human-level control through deep reinforcement learning.  
*Nature*, 518(7540):529–533.



Nikola Mrkšić, Diarmuid Ó Séaghdha, Tsung-Hsien Wen, Blaise Thomson, and Steve Young. 2017.

Neural belief tracker: Data-driven dialogue state tracking.  
In *Proceedings of ACL*, pages 1777–1788.



Nikola Mrkšić, Diarmuid Ó Séaghdha, Blaise Thomson, Milica Gašić, Pei-Hao Su, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2015.

Multi-domain dialog state tracking using recurrent neural networks.  
In *Proceedings of ACL*, pages 794–799.



Nikola Mrkšić, Ivan Vulić, Diarmuid Ó Séaghdha, Ira Leviant, Roi Reichart, Milica Gašić, Anna Korhonen, and Steve Young. 2017.

Semantic specialisation of distributional word vector spaces using monolingual and cross-lingual constraints.  
*Transactions of the ACL*, 5:309–324.

# (A Selection of) References IX



Jekaterina Novikova, Ondrej Dusek, Amanda Cercas Curry, and Verena Rieser. 2017.

Why we need new evaluation metrics for NLG.

*CoRR*, [abs/1707.06875](#).



Alice Oh and Alexander I. Rudnicky. 2002.

Stochastic natural language generation for spoken dialog systems.

*Computer Speech & Language*, 16(3-4):387–407.



Abhinav Rastogi, Dilek Hakkani-Tür, and Larry P. Heck. 2017.

Scalable multi-domain dialogue state tracking.

In *Proceedings of ASRU*, pages 561–568.



Iulian Vlad Serban, Alessandro Sordoni, Ryan Lowe, Laurent Charlin, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio. 2017.

A hierarchical latent variable encoder-decoder model for generating dialogues.

In *Proceedings of AAAI*, pages 3295–3301.



Pararth Shah, Dilek Hakkani-Tür, Gökhan Tür, Abhinav Rastogi, Ankur Bapna, Neha Nayak, and Larry P. Heck. 2018.

Building a conversational agent overnight with dialogue self-play.

# (A Selection of) References X



Hongjie Shi, Takashi Ushio, Mitsuru Endo, Katsuyoshi Yamagami, and Noriaki Horii. 2017.

A multichannel convolutional neural network for cross-language dialog state tracking.

*CoRR*, abs/1701.06247.



Pei-Hao Su, Paweł Budzianowski, Stefan Ultes, Milica Gasic, and Steve Young. 2017.

Sample-efficient actor-critic reinforcement learning with supervised data for dialogue management.

In *Proceedings of SIGDIAL*, pages 147–157.



Pei-Hao Su, Milica Gasic, Nikola Mrkšić, Lina M. Rojas Barahona, Stefan Ultes, David Vandyke, Tsung-Hsien Wen, and Steve Young. 2016.

On-line active reward learning for policy optimisation in spoken dialogue systems.

In *Proceedings of ACL*, pages 2431–2441.



Pei-Hao Su, David Vandyke, Milica Gašić, Nikola Mrkšić, Tsung-Hsien Wen, and Steve Young. 2015.

Reward shaping with recurrent neural networks for speeding up on-line policy learning in spoken dialogue systems.

In *Proceedings of SIGDIAL*, pages 417–421.

# (A Selection of) References XI



Blaise Thomson, Milica Gasic, Matthew Henderson, Pirros Tsiakoulis, and Steve J. Young. 2012.

N-best error simulation for training spoken dialogue systems.

In *Proceedings of SLT*, pages 37–42.



Blaise Thomson and Steve Young. 2010.

Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems.

*Computer Speech and Language*.



Van-Khanh Tran and Le-Minh Nguyen. 2017.

Natural language generation for spoken dialogue system using RNN encoder-decoder networks.

In *Proceedings of CoNLL*, pages 442–451.



Stefan Ultes, Lina Maria Rojas-Barahona, Pei-Hao Su, David Vandyke, Dongho Kim, Iñigo Casanueva, Pawel Budzianowski, Nikola Mrkšić, Tsung-Hsien Wen, Milica Gašić, and Steve J. Young. 2017.

Pydial: A multi-domain statistical dialogue system toolkit.

In *Proceedings of ACL Demos*, pages 73–78.

# (A Selection of) References XII



David Vandyke, Pei-hao Su, Milica Gasic, Nikola Mrksic, Tsung-Hsien Wen, and Steve J. Young. 2015.

Multi-domain dialogue success classifiers for policy training.

In *Proceedings of ASRU*, pages 763–770.



Oriol Vinyals and Quoc V. Le. 2015.

A neural conversational model.

*CoRR*, abs/1506.05869.



Ivan Vulić, Goran Glavaš, Nikola Mrkšić, and Anna Korhonen. 2018.

Post-specialisation: Retrofitting vectors of words unseen in lexical resources.

In *Proceedings of NAACL-HLT*.



Ivan Vulić, Nikola Mrkšić, Roi Reichart, Diarmuid Ó Séaghdha, Steve Young, and Anna Korhonen. 2017.

Morph-fitting: Fine-tuning word vector spaces with simple language-specific rules.

In *Proceedings of ACL*, pages 56–68.



Marilyn A. Walker, Owen Rambow, and Monica Rogati. 2002.

Training a sentence planner for spoken dialogue using boosting.

*Computer Speech & Language*, 16(3-4):409–433.



# (A Selection of) References XIII



Joseph Weizenbaum. 1966.  
ELIZA - A computer program for the study of natural language communication between man and machine.  
*Communications of the ACM*, 9(1):36–45.



Tsung-Hsien Wen, Milica Gasic, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. 2015a.  
Semantically conditioned LSTM-based natural language generation for spoken dialogue systems.  
In *Proceedings of EMNLP*, pages 1711–1721.



Tsung-Hsien Wen, Milica Gašić, Dongho Kim, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. 2015b.  
Stochastic language generation in dialogue using recurrent neural networks with convolutional sentence reranking.  
In *Proceedings of SIGDIAL*, pages 275–284.



Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Lina M. Rojas-Barahona, Pei-Hao Su, David Vandyke, and Steve Young. 2016.  
Multi-domain neural network language generation for spoken dialogue systems.  
In *Proceedings of NAACL-HLT*, pages 120–129.

# (A Selection of) References XIV



Tsung-Hsien Wen, Yishu Miao, Phil Blunsom, and Steve J. Young. 2017a.  
Latent intention dialogue models.  
*In Proceedings of ICML*, pages 3732–3741.



Tsung-Hsien Wen, David Vandyke, Nikola Mrkšić, Milica Gasic, Lina M. Rojas Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. 2017b.  
A network-based end-to-end trainable task-oriented dialogue system.  
*In Proceedings of EACL*, pages 438–449.



Jason D. Williams, Kavosh Asadi, and Geoffrey Zweig. 2017.  
Hybrid code networks: Practical and efficient end-to-end dialog control with supervised and reinforcement learning.  
*In Proceedings of ACL*, pages 665–677.



Jason D. Williams, Antoine Raux, and Matthew Henderson. 2016.  
The Dialog State Tracking Challenge Series: A review.  
*Dialogue & Discourse*, 7(3):4–33.



Jason D. Williams and Steve Young. 2007.  
Partially observable markov decision processes for spoken dialog systems.  
*Computer Speech and Language*, 21:393–422.

## (A Selection of) References XV



Jason D. Williams and Geoffrey Zweig. 2016.  
End-to-end LSTM-based dialog control optimized with supervised and reinforcement learning.  
*CoRR*, abs/1606.01269.



Xuesong Yang, Yun-Nung Chen, Dilek Z. Hakkani-Tür, Paul Crook, Xiujuan Li, Jianfeng Gao, and Li Deng. 2017.  
End-to-end joint learning of natural language understanding and dialogue manager.  
In *Proceedings of ICASSP*, pages 5690–5694.



Tiancheng Zhao and Maxine Eskenazi. 2016.  
Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning.  
In *Proceedings of SIGDIAL*, pages 1–10.

תודה  
Dankie Gracias  
Спасибо شُكراً  
Merci Takk  
Köszönjük Terima kasih  
Grazie Dziękujemy Děkojame  
Ďakujeme Vielen Dank Paldies  
Kiitos Täname teid 谢谢  
**Thank You** Tak  
感謝您 Obrigado Teşekkür Ederiz  
Σας Ευχαριστούμ 감사합니다  
ขอบคุณ  
Bedankt Děkujeme vám  
ありがとうございます  
Tack

Many many thanks to Tsung-Hsien Wen, Vivian Chen, Matthew Henderson, Dilek Hakkani-Tür, Asli Celikyilmaz, Parath Shah, Abhinav Rastogi, Milica Gašić, Lihong Li, Ankur Bapna, Gokhan Tür, and others for sharing their amazing slides!