OPEN DATA SCIENCE CONFERENCE

Boston | May 1 - 4 2018

#ODSC

@ODSC

Democratise Conversational AI

Scaling Academic Research to Industrial Applications

Shawn (Tsung-Hsien) Wen
shawnwen@poly-ai.com, www.poly-ai.com

POLYAi

Conversational AI in Fiction

Evolution of Conversational AI

Turing Imitation Game, 1950

The Dartmouth proposal, 1955

The ELIZA chatbot, 1964

World Wide Web - largest knowledge base, 1989

AT&T MATCH Multimodal System, 2002

Launch of Apple Siri, 2011

Virtual Assistant platforms, 2017

Future of CAIs
- Integration to many services
- Multimodal
- Multilingual
- Multi-turn
- Ubiquitous

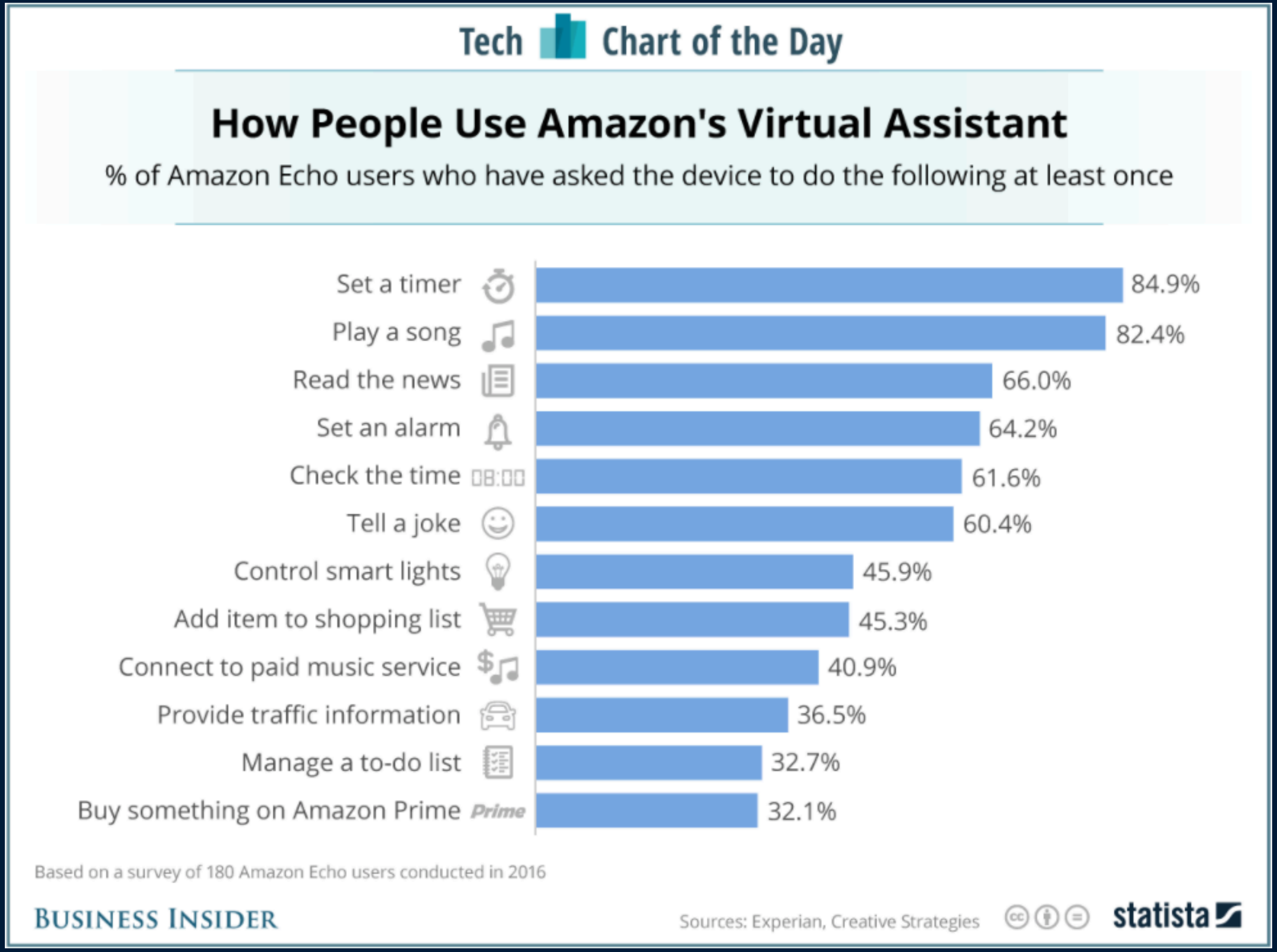
A Not-So-Far-Away Future

"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."

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

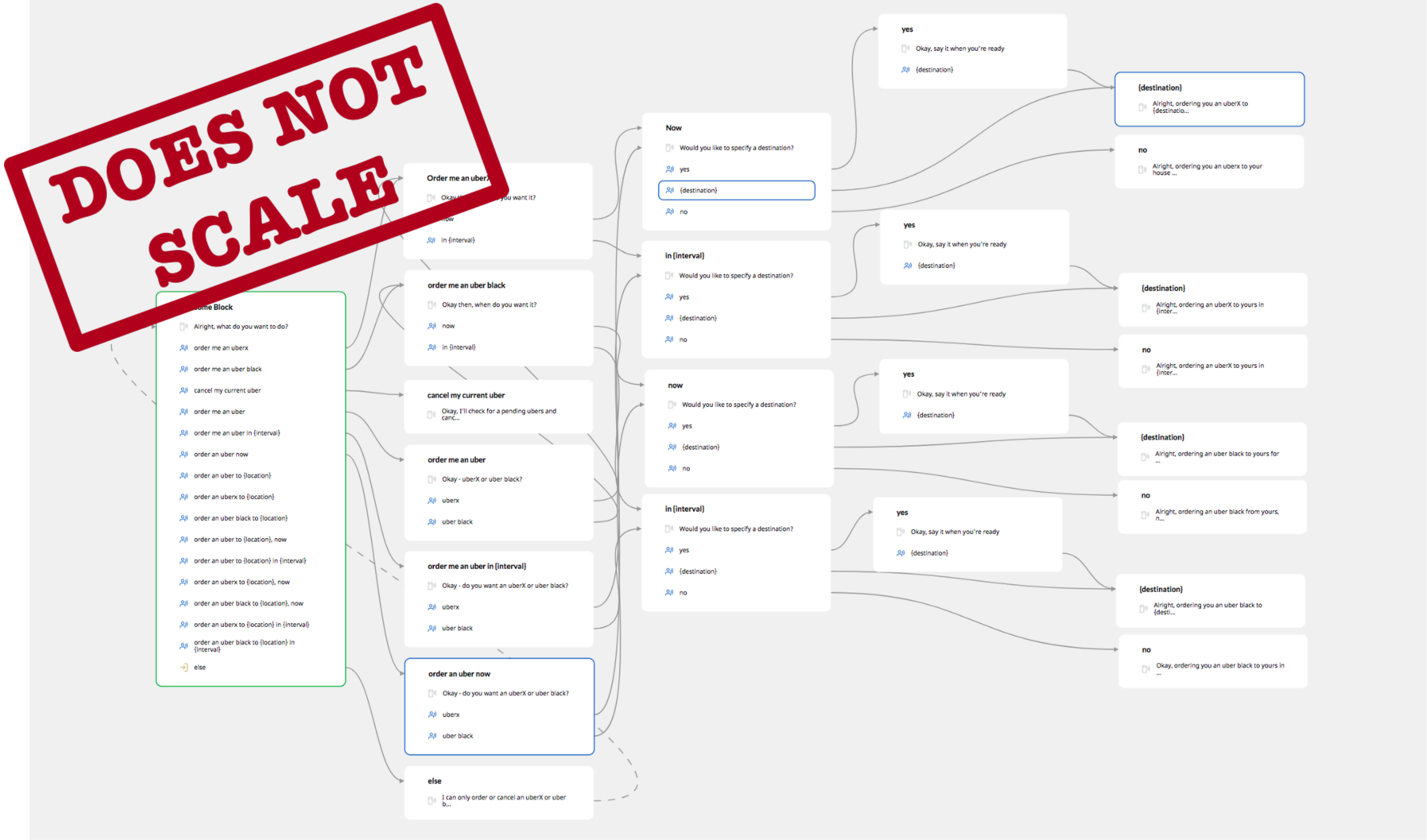
statista

POLYAI

Conversational AI in 2018

- Virtual Assistants are currently used for **simple tasks** - the more complex (and useful) ones are implemented and supported by the platform providers themselves
- Third-party service providers are finding it **hard to connect** their services to existing virtual assistant platforms
  - This major **technical block** stands in the way of virtual assistants becoming the next iPhone - they are waiting for their AppStore to come to life

Current Industry Practice - Flowcharts



DOES NOT SCALE

ome Block
Alright, what do you want to do?
order me an uberx
order me an uber black
cancel my current uber
order me an uber
order me an uber in {interval}
order an uber now
order an uber to {location}
order an uberx to {location}
order an uber black to {location}
order an uber to {location}, now
order an uber to {location} in {interval}
order an uberx to {location}, now
order an uber black to {location}, now
order an uberx to {location} in {interval}
order an uber black to {location} in {interval}
else

Order me an uberX
Okay you want it?
ow
in {interval}

order me an uber black
Okay then, when do you want it?
now
in {interval}

cancel my current uber
Okay, I'll check for a pending ubers and canc...

order me an uber
Okay - uberX or uber black?
uberx
uber black

order me an uber in {interval}
Okay - do you want an uberX or uber black?
uberx
uber black

order an uber now
Okay - do you want an uberX or uber black?
uberx
uber black

else
I can only order or cancel an uberX or uber b...

Now
Would you like to specify a destination?
yes
{destination}
no

in {interval}
Would you like to specify a destination?
yes
{destination}
no

now
Would you like to specify a destination?
yes
{destination}
no

in {interval}
Would you like to specify a destination?
yes
{destination}
no

yes
Okay, say it when you're ready
{destination}

yes
Okay, say it when you're ready
{destination}

yes
Okay, say it when you're ready
{destination}

yes
Okay, say it when you're ready
{destination}

{destination}
Alright, ordering you an uberX to {destinatio...

no
Alright, ordering you an uberx to your house ...

{destination}
Alright, ordering an uberX to yours in {inter...

no
Alright, ordering an uberX to yours in {inter...

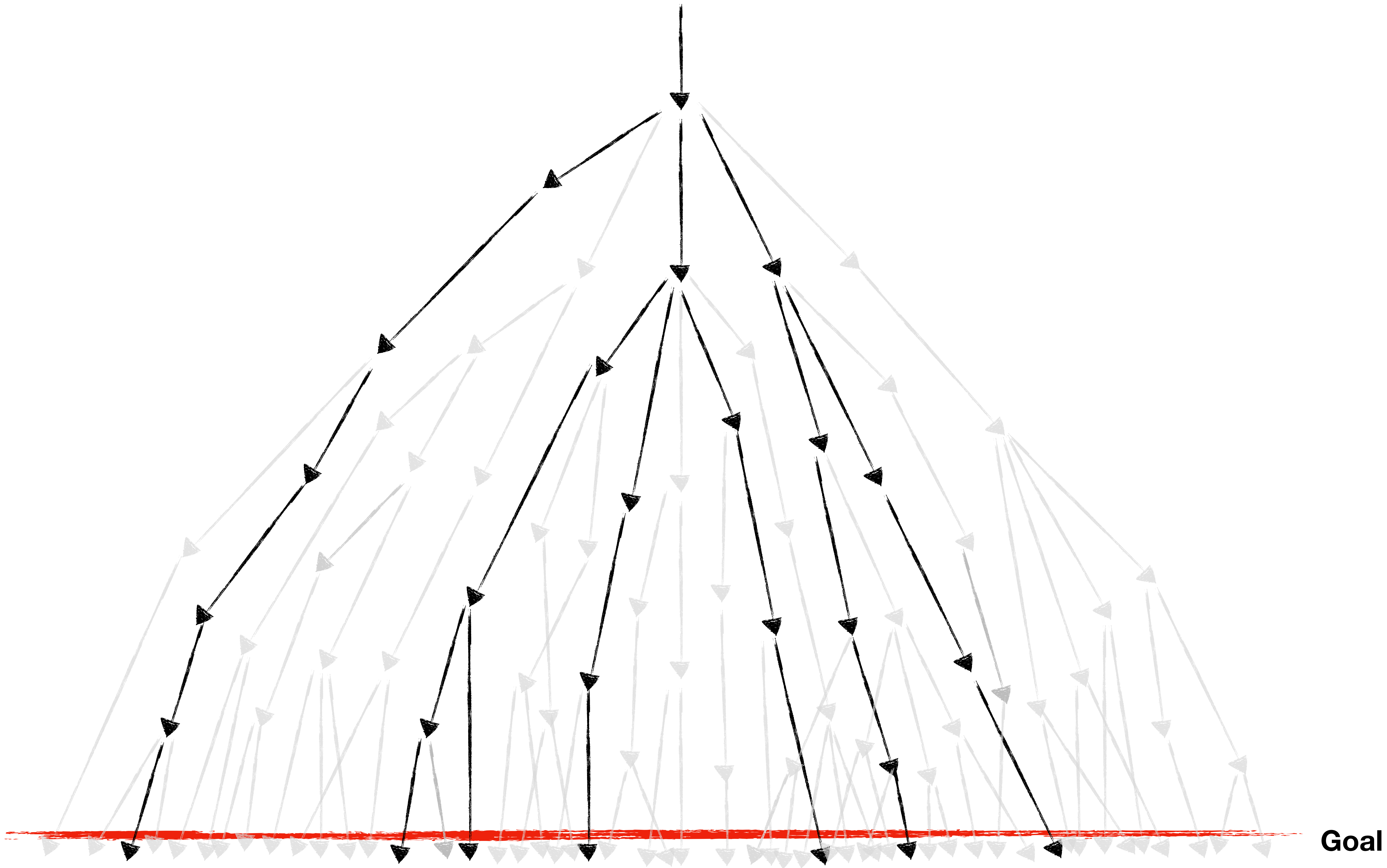
{destination}
Alright, ordering an uber black to yours for ...

no
Alright, ordering an uber black from yours, n...

{destination}
Alright, ordering you an uber black to {desti...

no
Okay, ordering you an uber black to yours in ...

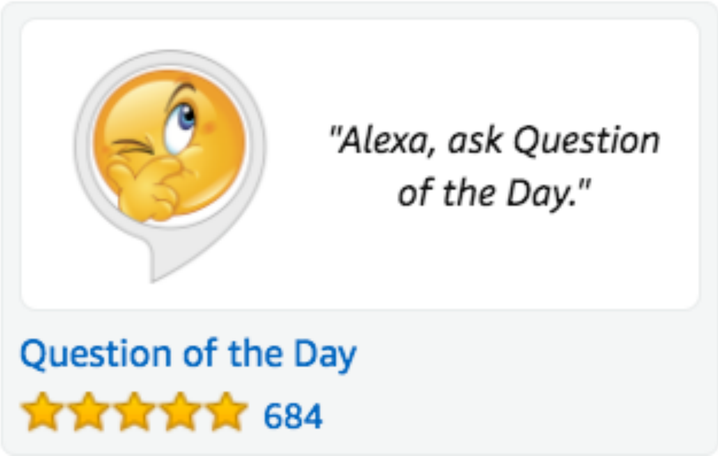
The Happy Path Approach



- Set an alarm for 8.30am.
- Alarm at 8.30am, please.
- Wake me up in seven hours.
- Alarm next morning at 8.30.
- Wake me up at 8.30am.
- An alarm for half eight.
- Turn on my 8.30 alarm.
- …

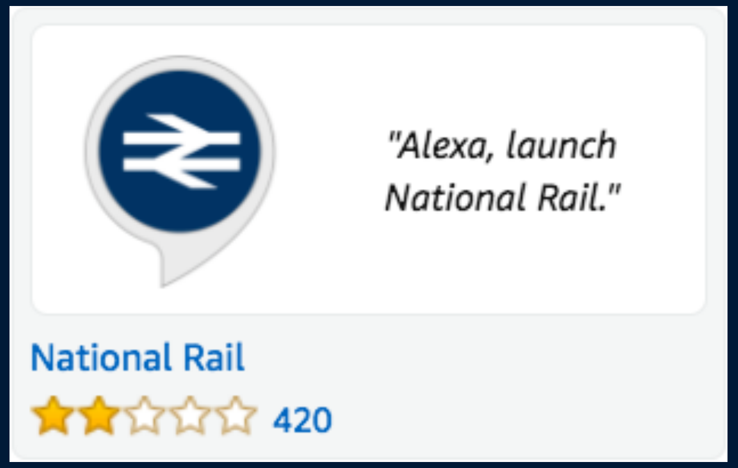
Simple Flowchart Tools = Simple Voice-Based Apps

Top Alexa Skills are still simple single-turn commands


"Alexa, ask Question of the Day."

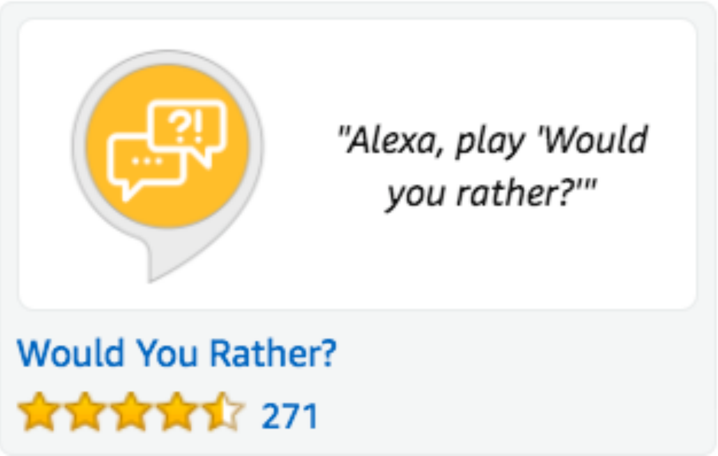
Question of the Day

★★★★★ 684


"Alexa, launch National Rail."

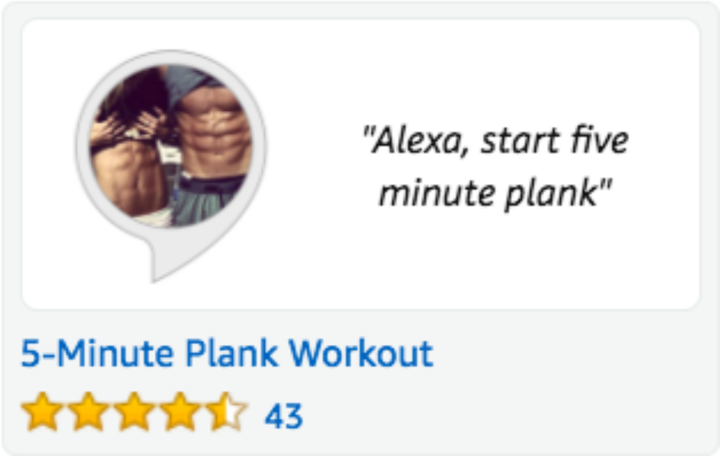
National Rail

★★☆☆☆ 420


"Alexa, play 'Would you rather?'"

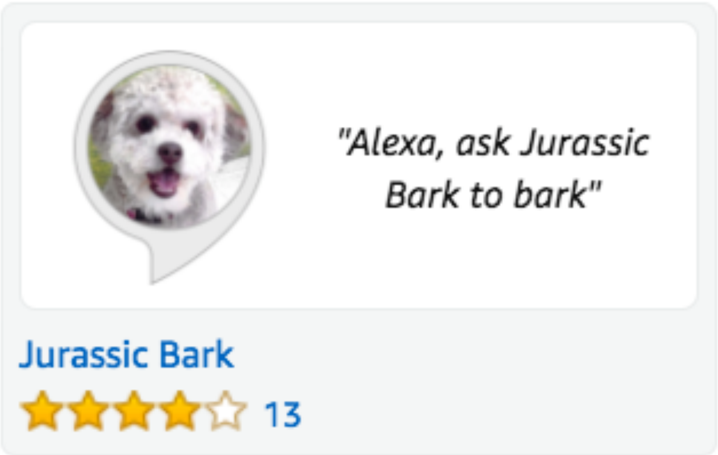
Would You Rather?

★★★★⯪ 271


"Alexa, start five minute plank"

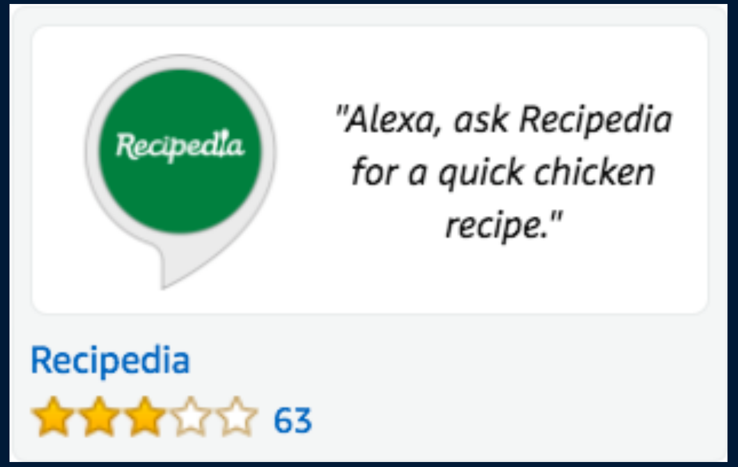
5-Minute Plank Workout

★★★★⯪ 43


"Alexa, ask Jurassic Bark to bark"

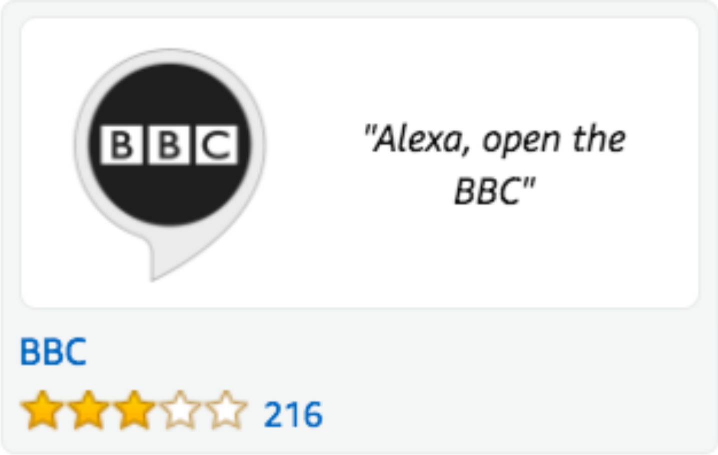
Jurassic Bark

★★★★☆ 13


"Alexa, ask Recipedia for a quick chicken recipe."

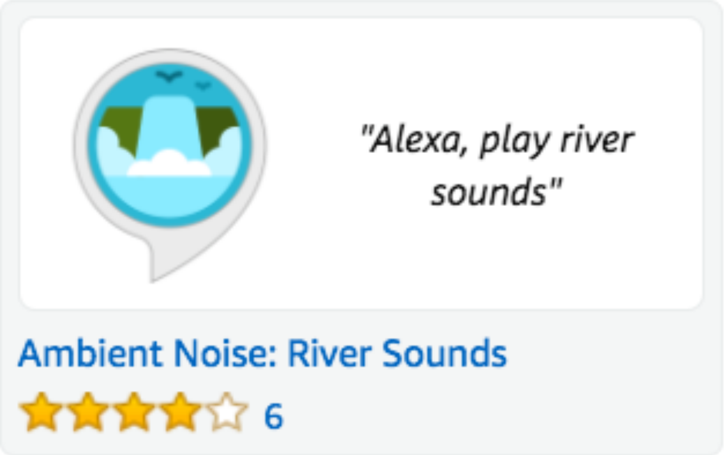
Recipedia

★★★☆☆ 63


"Alexa, open the BBC"

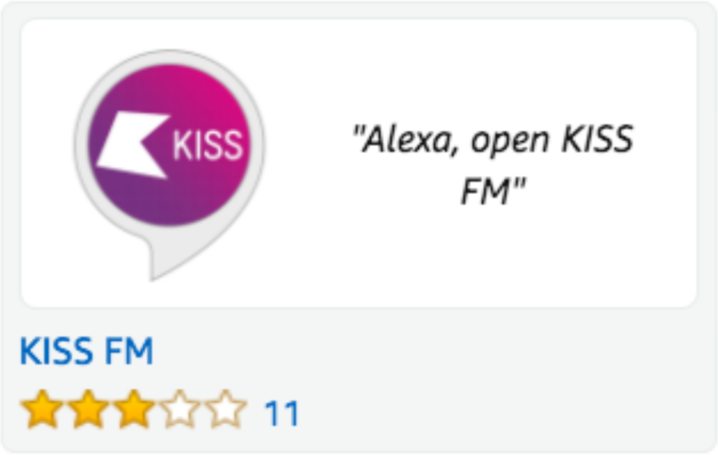
BBC

★★★☆☆ 216


"Alexa, play river sounds"

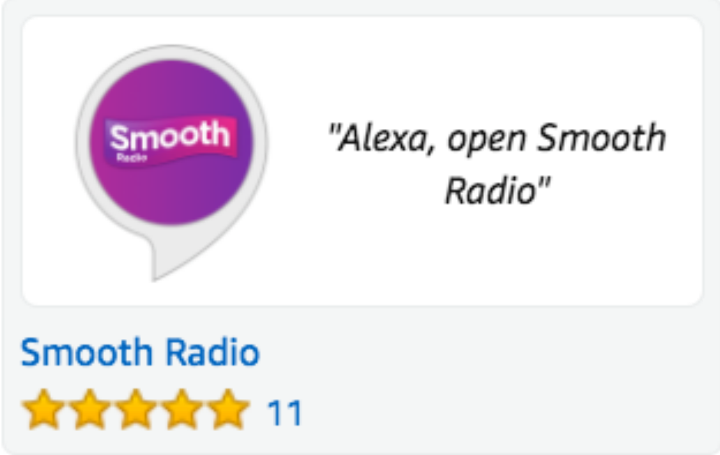
Ambient Noise: River Sounds

★★★★☆ 6


"Alexa, open KISS FM"

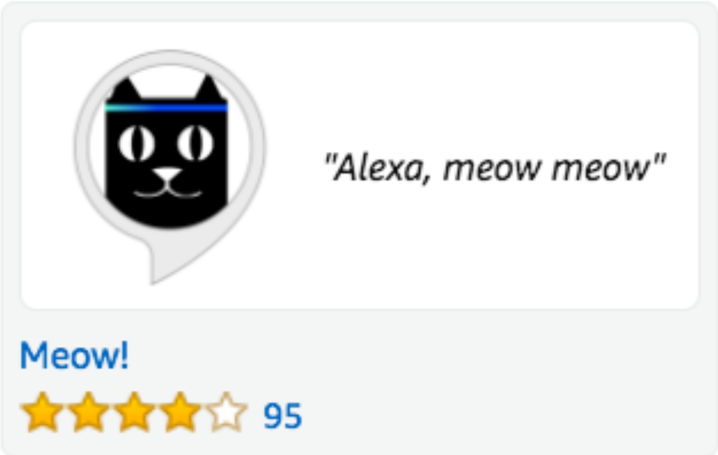
KISS FM

★★★☆☆ 11


"Alexa, open Smooth Radio"

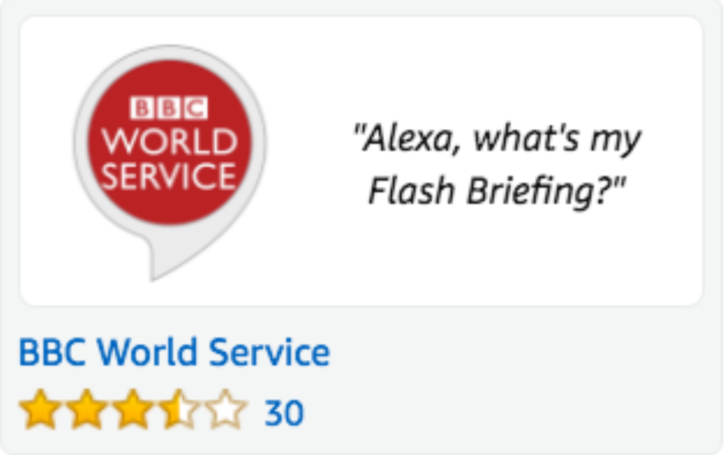
Smooth Radio

★★★★★ 11


"Alexa, meow meow"

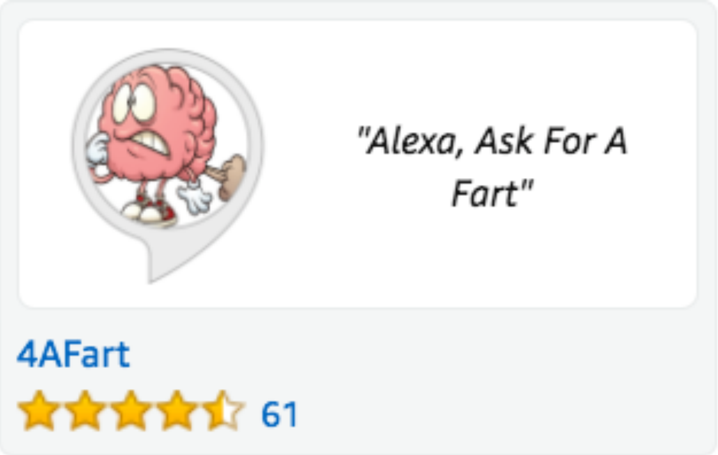
Meow!

★★★★☆ 95


"Alexa, what's my Flash Briefing?"

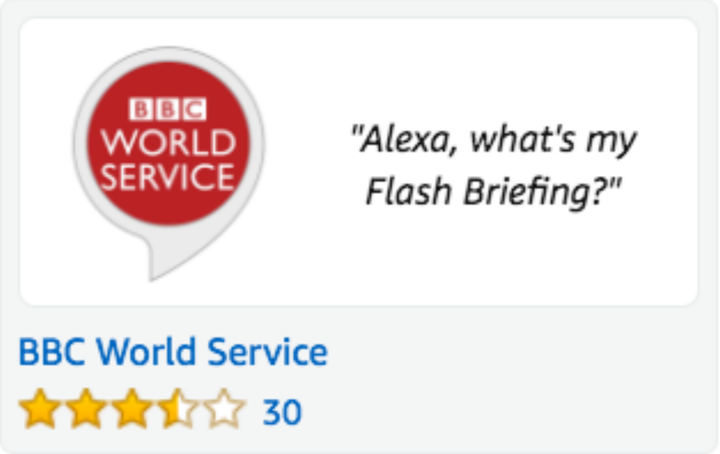
BBC World Service

★★★⯪☆ 30


"Alexa, Ask For A Fart"

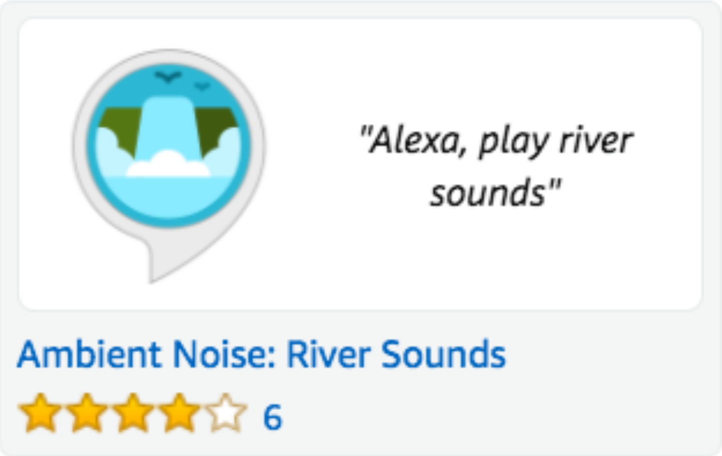
4AFart

★★★★⯪ 61


"Alexa, what's my Flash Briefing?"

BBC World Service

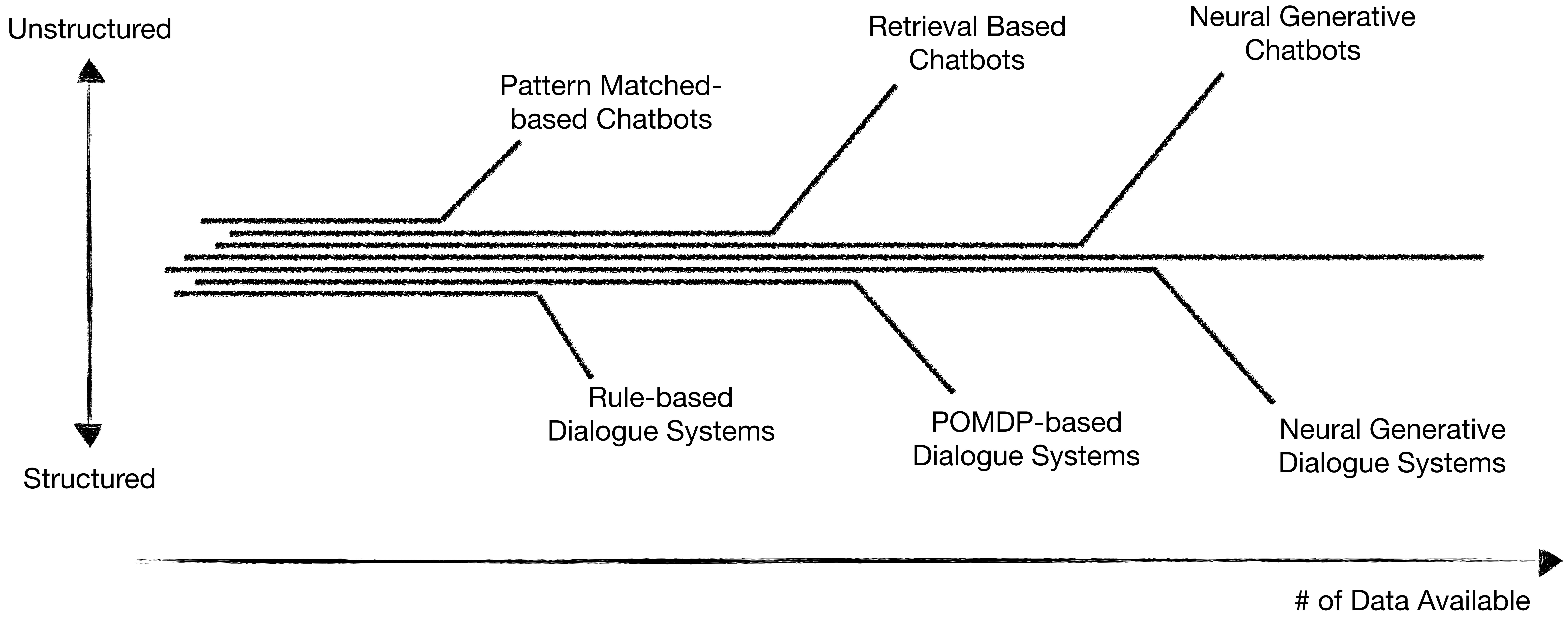
★★★⯪☆ 30


"Alexa, play river sounds"

Ambient Noise: River Sounds

★★★★☆ 6

Conversational AI Approaches



Unstructured

Structured

Pattern Matched-based Chatbots

Retrieval Based Chatbots

Neural Generative Chatbots

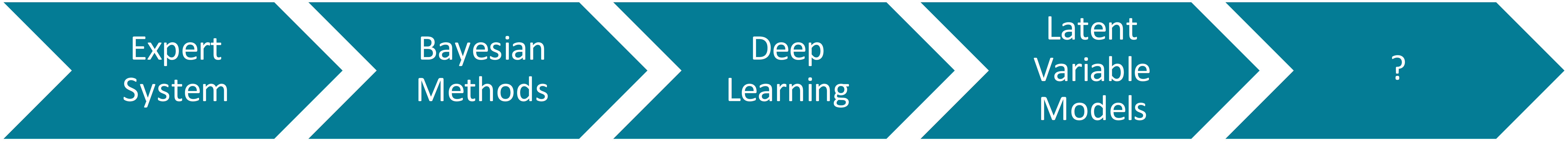
Rule-based Dialogue Systems

POMDP-based Dialogue Systems

Neural Generative Dialogue Systems

# of Data Available

A Single-Direction Development?



Expert System

Bayesian Methods

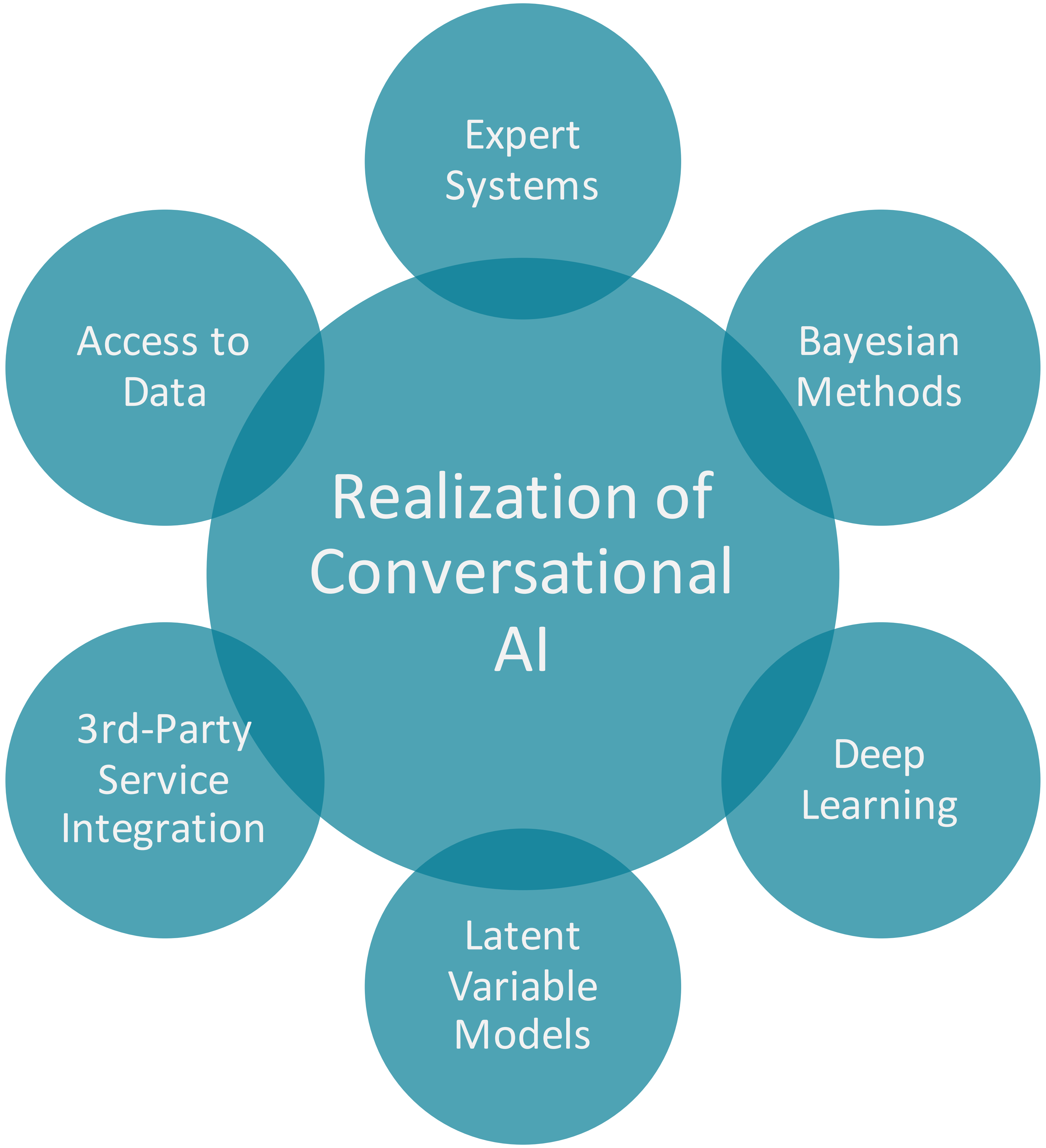
Deep Learning

Latent Variable Models

?

"Is new technology always better than what we have already had?"

A Collective Effort!



Realizing Conversational AI requires efforts from multiple parties:

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

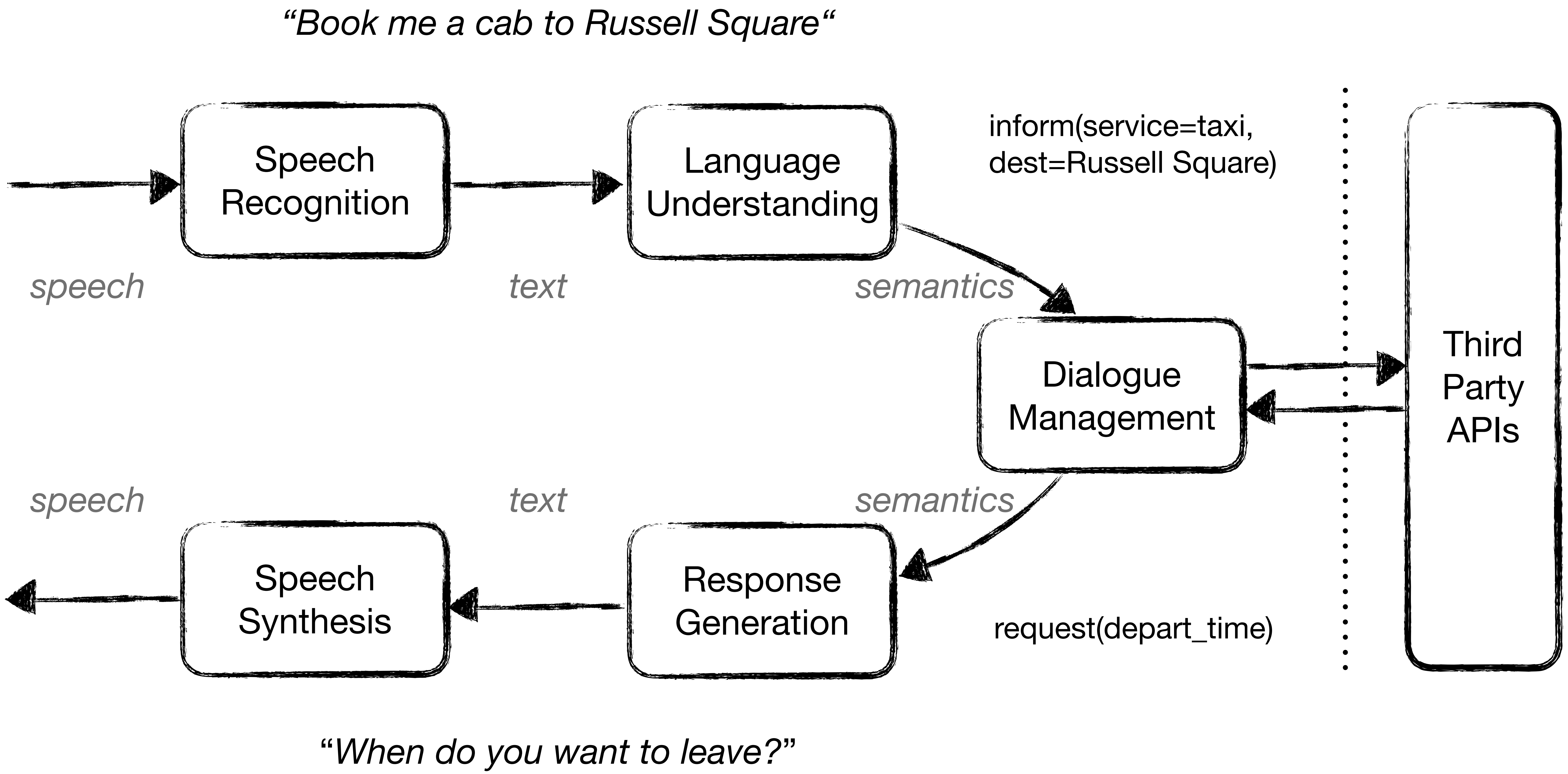
Realizing Conversational AI also requires a hybrid of approaches to tackle different parts of the pipeline:

- 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.

Outline

- Conversational AI Architecture
- Overview of Conversational AI Research
    - End-to-End Chatbots
    - Learning-based Goal-oriented Systems
- Data! Data! Data!
- Conclusion

Conversational AI - Architecture



“Book me a cab to Russell Square“

Speech Recognition

Language Understanding

inform(service=taxi,
dest=Russell Square)

speech

text

semantics

Dialogue Management

Third Party APIs

speech

text

semantics

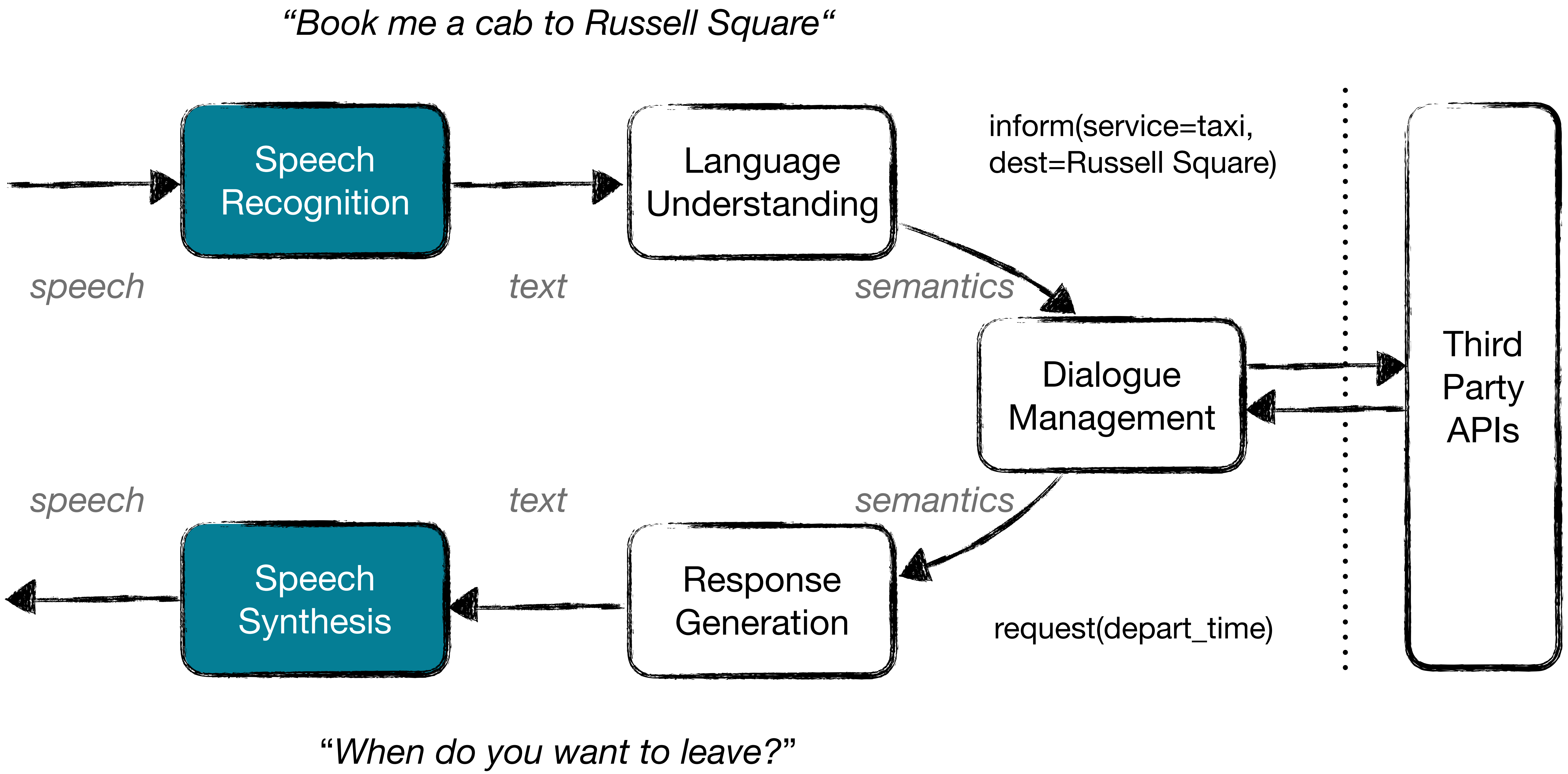
Speech Synthesis

Response Generation

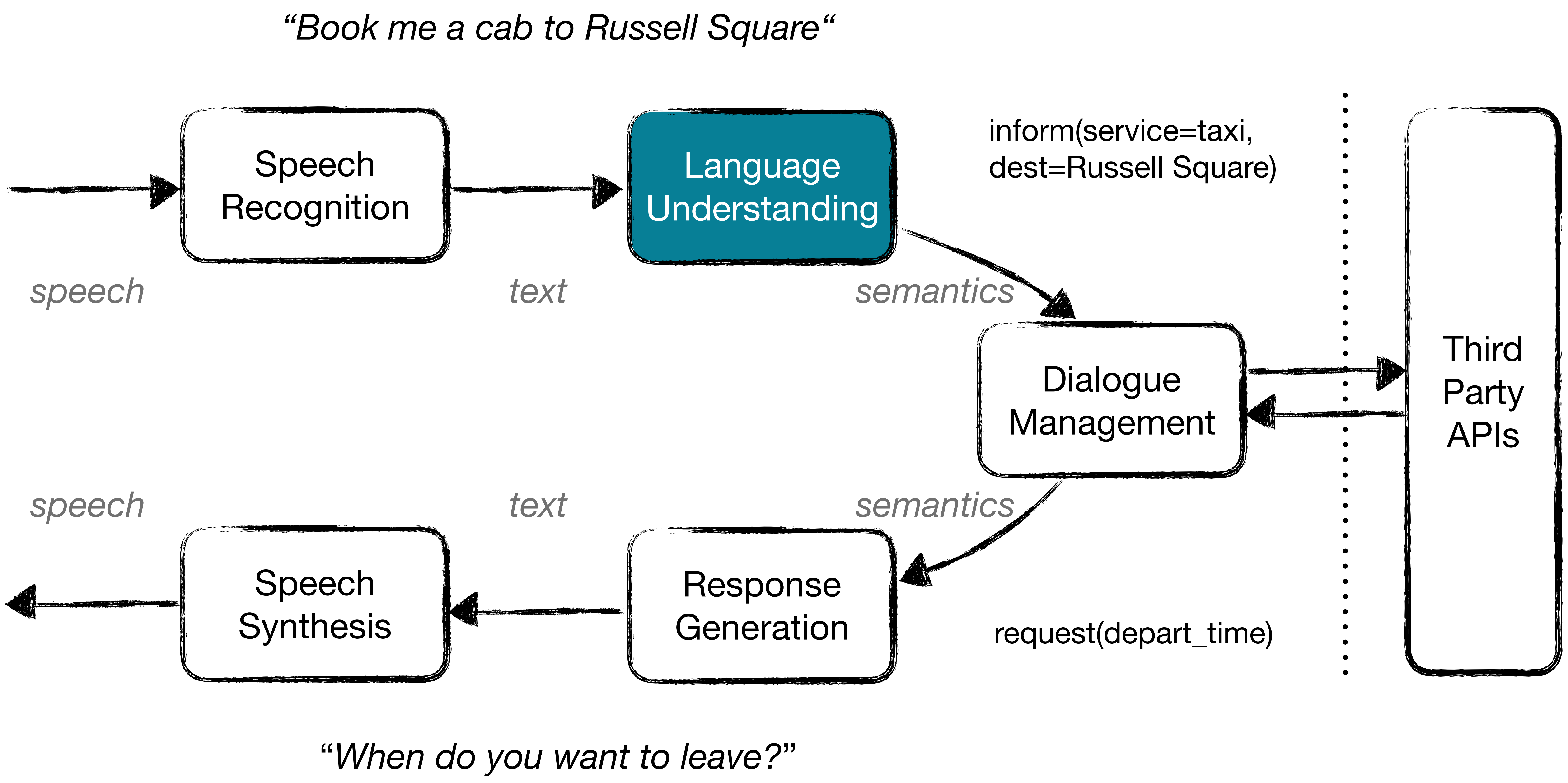
request(depart_time)

“When do you want to leave?”

Conversational AI - Architecture



Conversational AI - Architecture



“Book me a cab to Russell Square“

Speech Recognition

Language Understanding

inform(service=taxi,
dest=Russell Square)

Dialogue Management

Third Party APIs

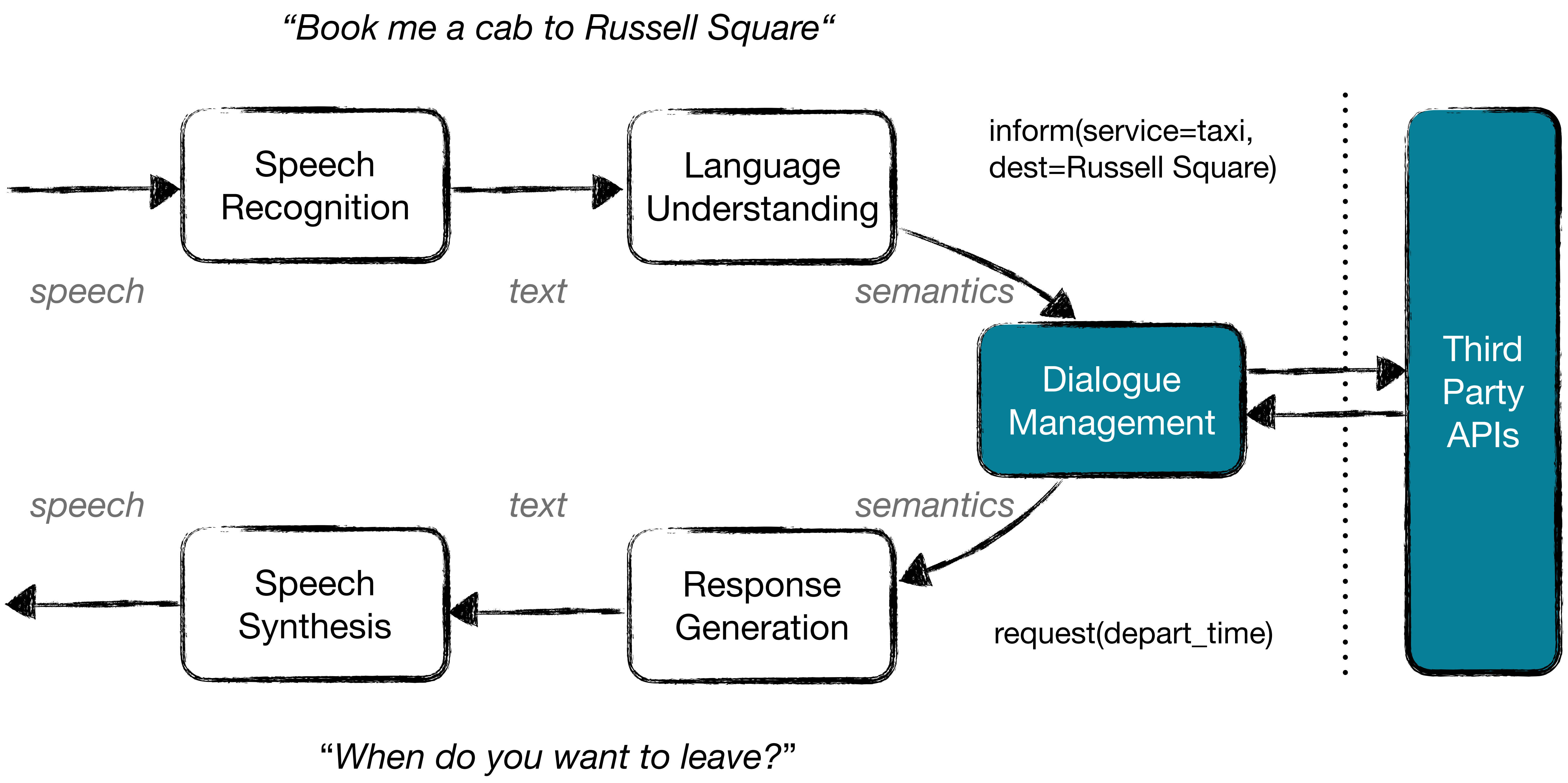
Response Generation

request(depart_time)

Speech Synthesis

“When do you want to leave?”

Conversational AI - Architecture



"Book me a cab to Russell Square"

Speech Recognition

Language Understanding

inform(service=taxi,
dest=Russell Square)

speech

text

semantics

Dialogue Management

Third Party APIs

speech

text

semantics

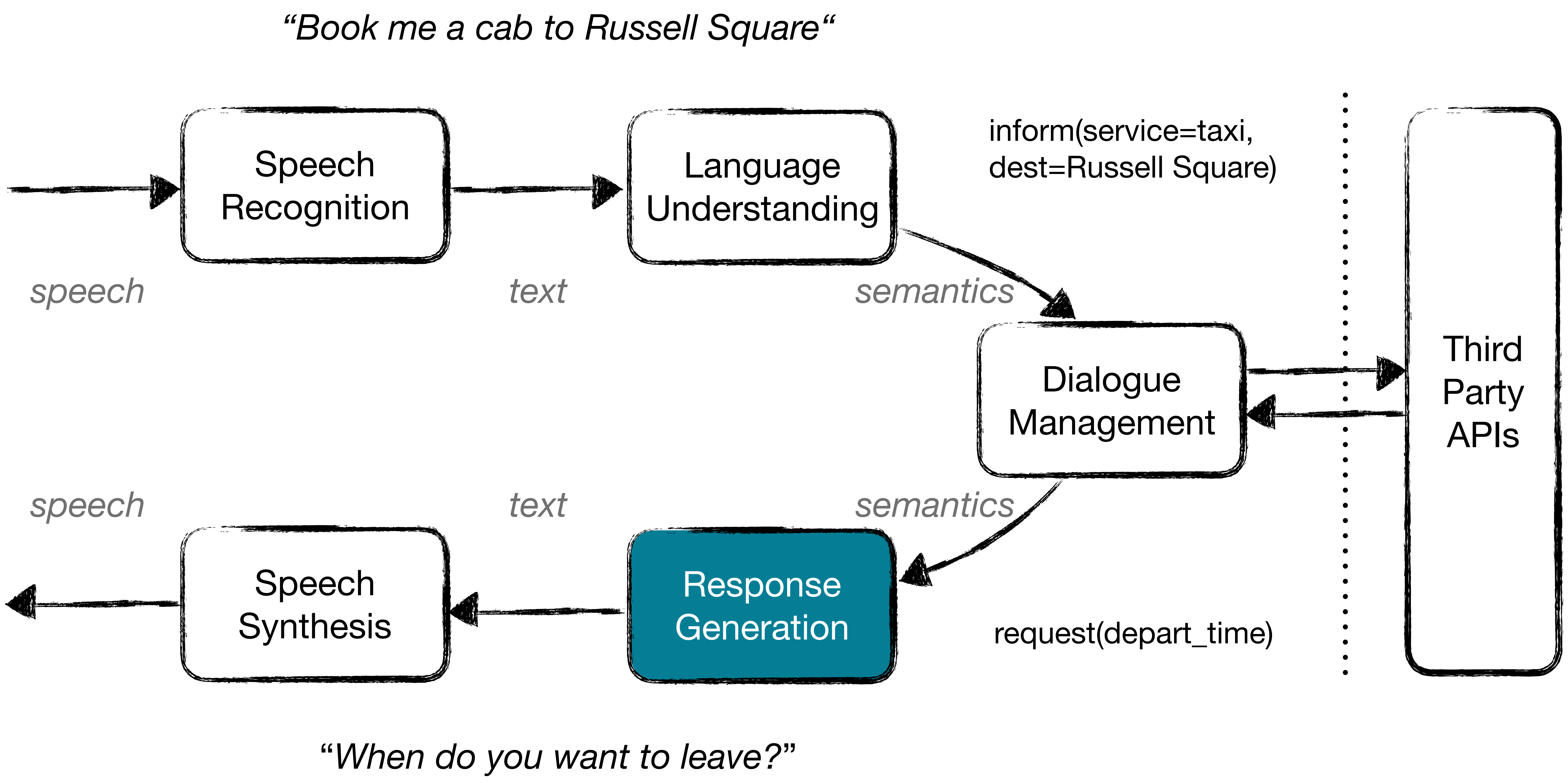
Speech Synthesis

Response Generation

request(depart_time)

"When do you want to leave?"

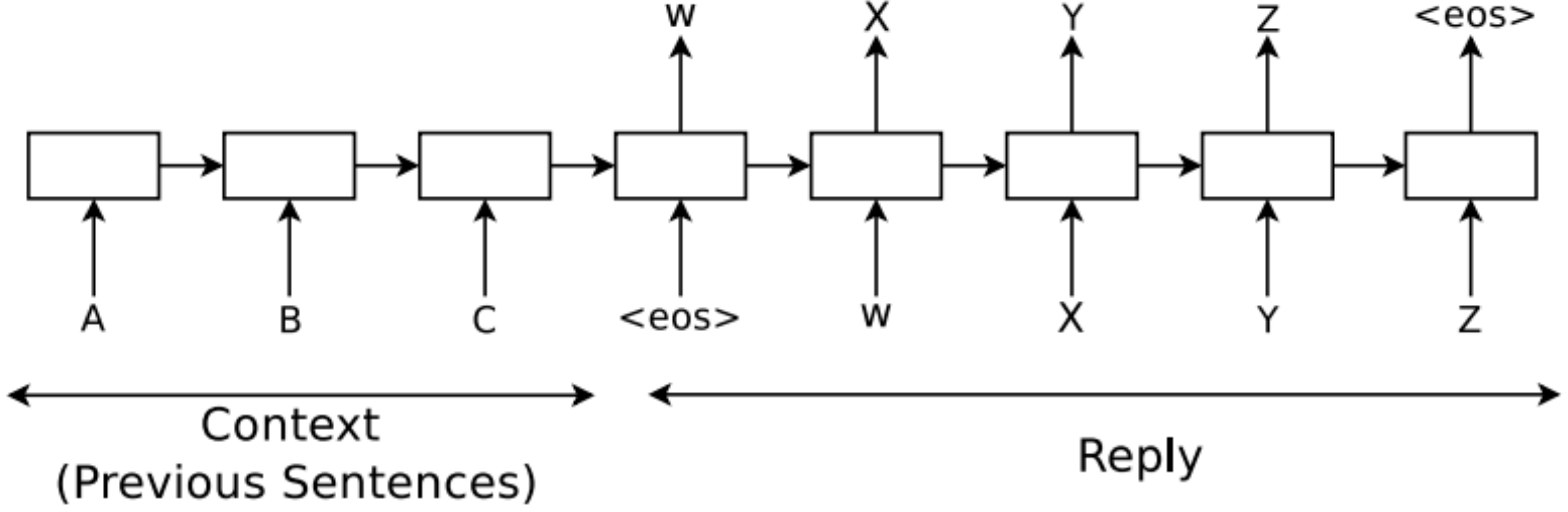
Conversational AI - Architecture

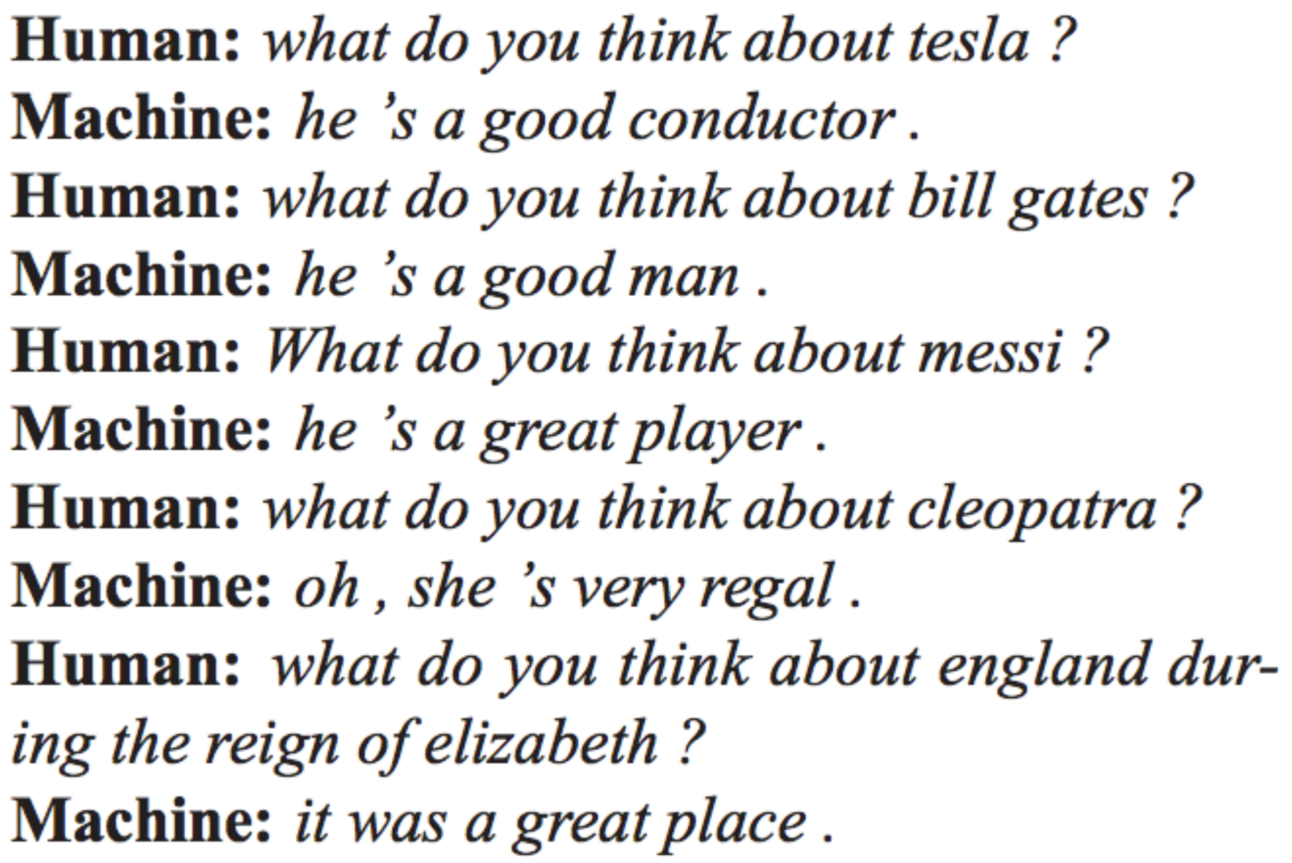


End-to-End Chatbots

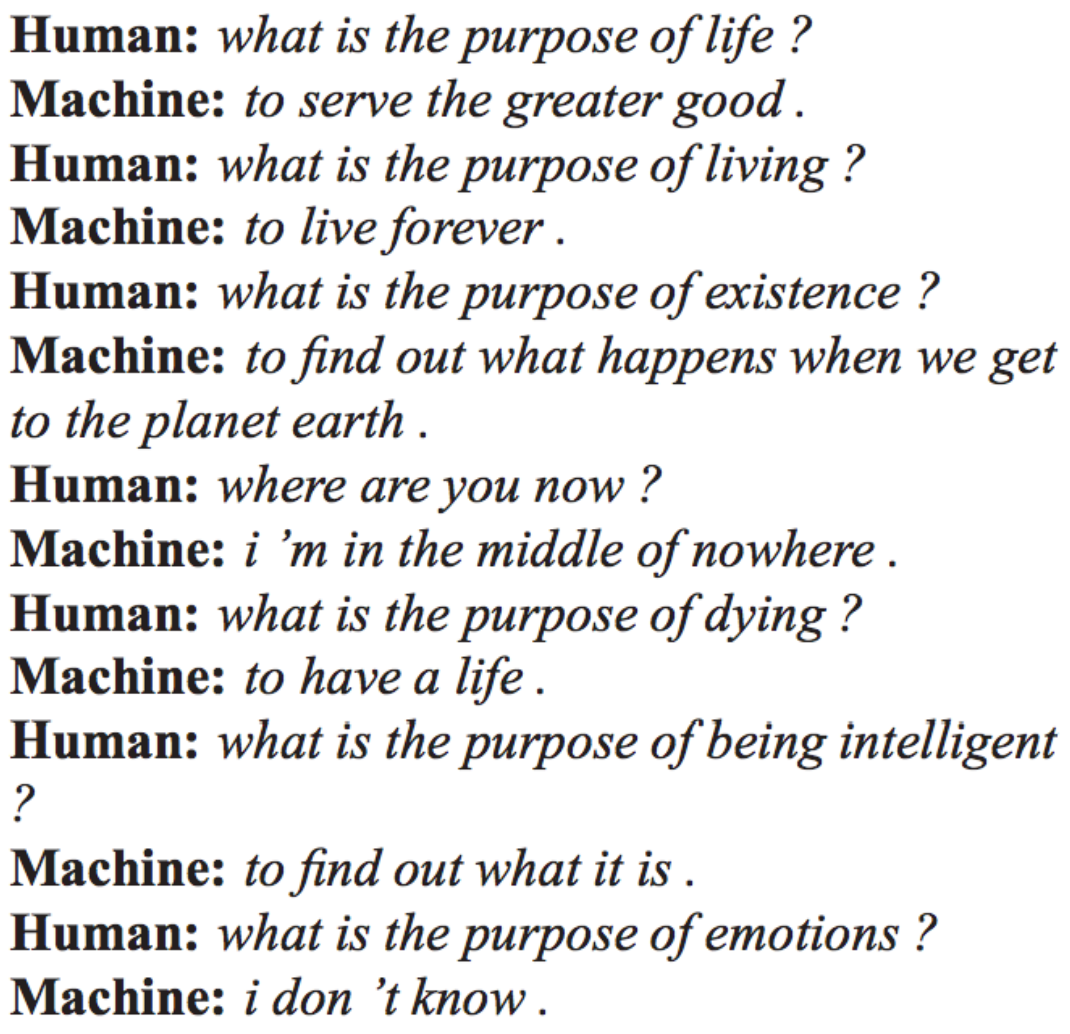
Sequence-to-Sequence Conversation Model

Vinyals & Le, 2015





Human: *what do you think about tesla ?*
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 .*

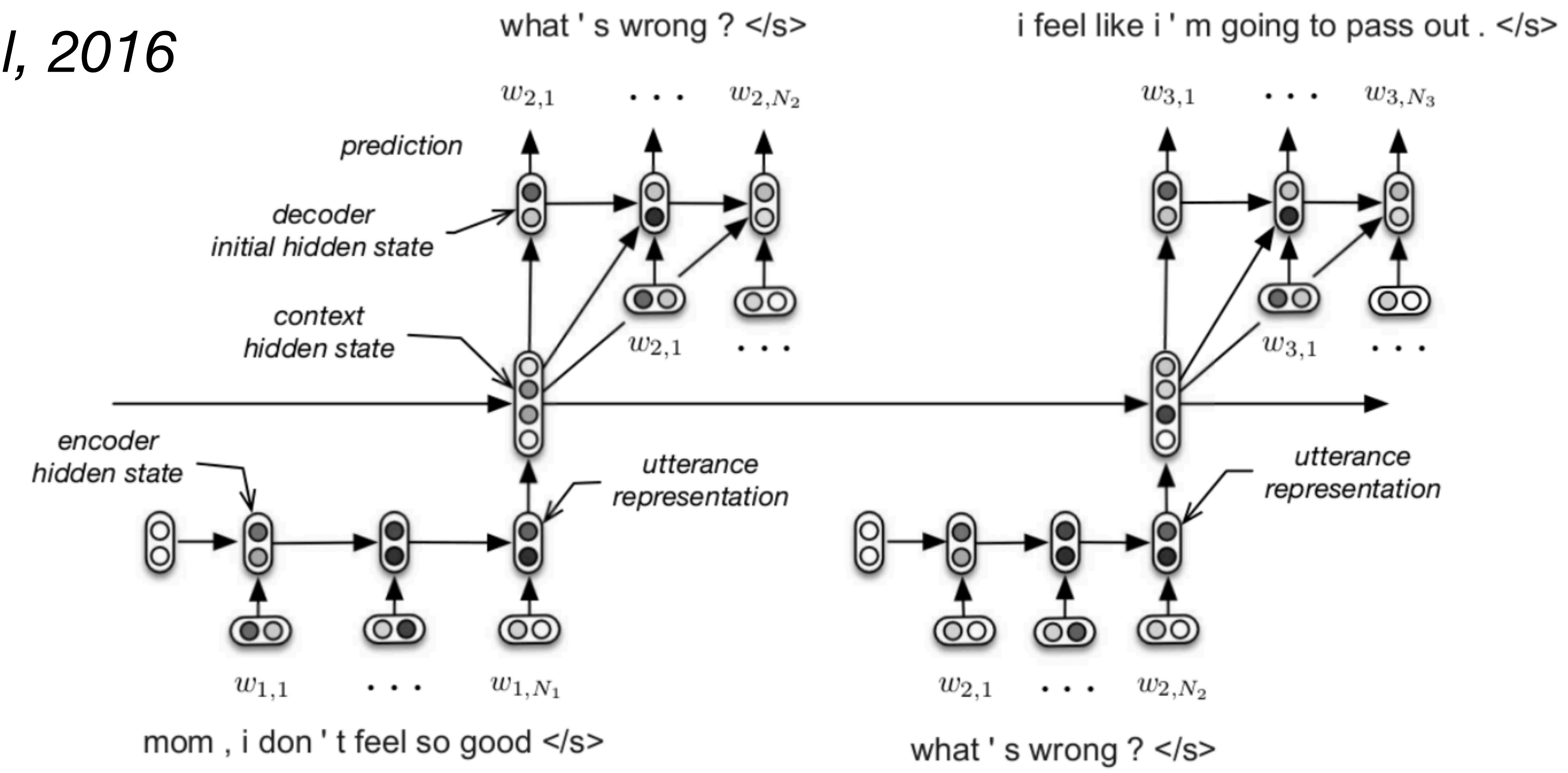


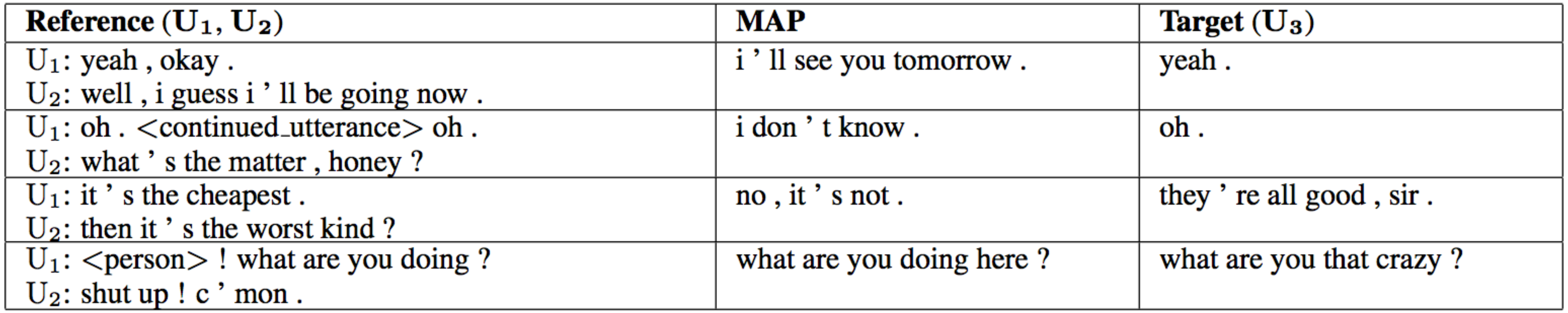
Human: *what is the purpose of life ?*
Machine: *to serve the greater good .*
Human: *what is the purpose of living ?*
Machine: *to live forever .*
Human: *what is the purpose of existence ?*
Machine: *to find out what happens when we get to the planet earth .*
Human: *where are you now ?*
Machine: *i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

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

Hierarchical RNN for Dialogue

Serban et al, 2016



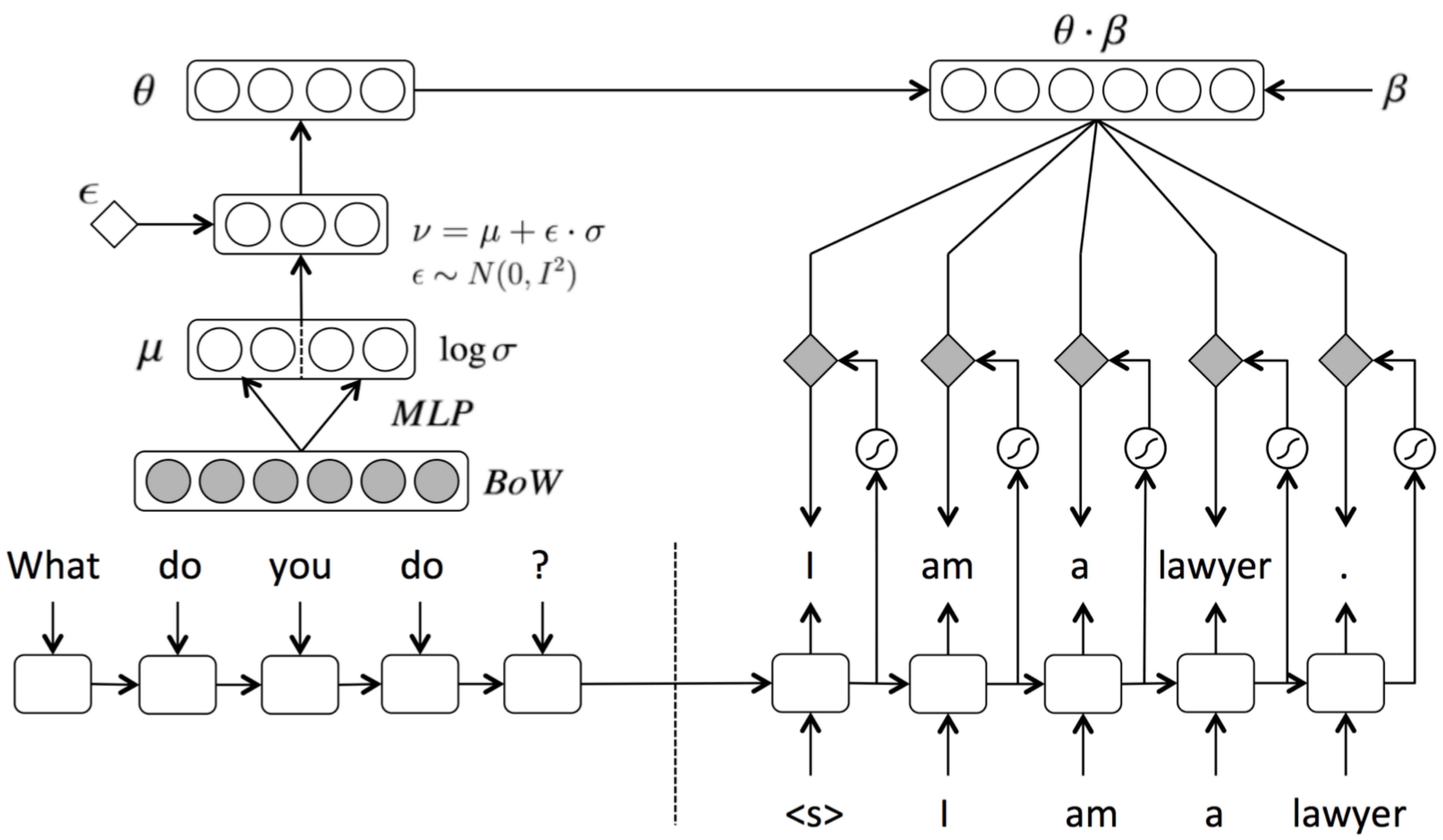


Reference (U_1, U_2)	**MAP**	**Target (U_3)**
U_1: yeah , okay . U_2: well , i guess i ' ll be going now .	i ' ll see you tomorrow .	yeah .
U_1: oh . <continued_utterance> oh . U_2: what ' s the matter , honey ?	i don ' t know .	oh .
U_1: it ' s the cheapest . U_2: then it ' s the worst kind ?	no , it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

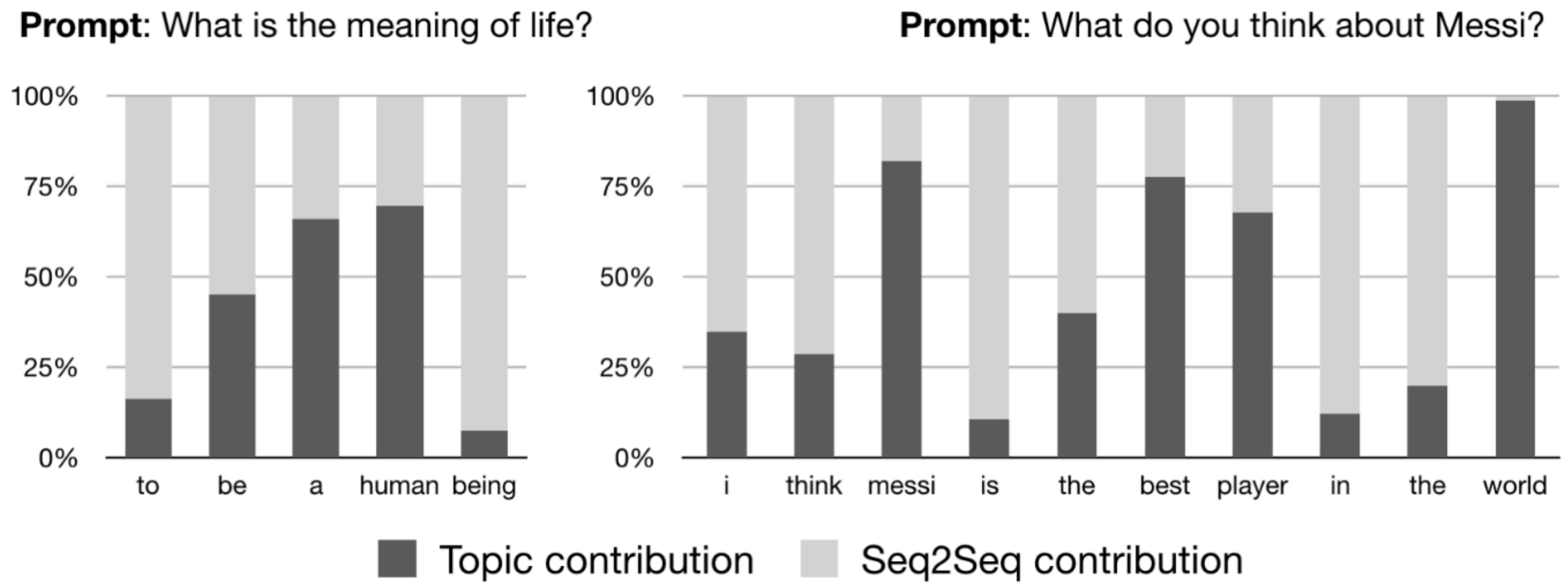
Problems of the Approach

- Problems
  - Generic, incoherent, and non-causal responses.
  - The goal is to “*response like a human being*”, not very useful for business use cases.
  - Lack of integration to knowledge base or 3rd party services.
  - Evaluation is super subjective and difficult.

An Improvement - Latent Topic Conversational Model



- Combine topic and seq2seq models.
- Optimise as a Variational Autoencoder.
- Use the topic component to control meaning, while adopt seq2seq to produce grammatical responses.



Prompt: What is the purpose of existence?
LTCM: To be able to see the world.
To be able to see things.
To make it better.

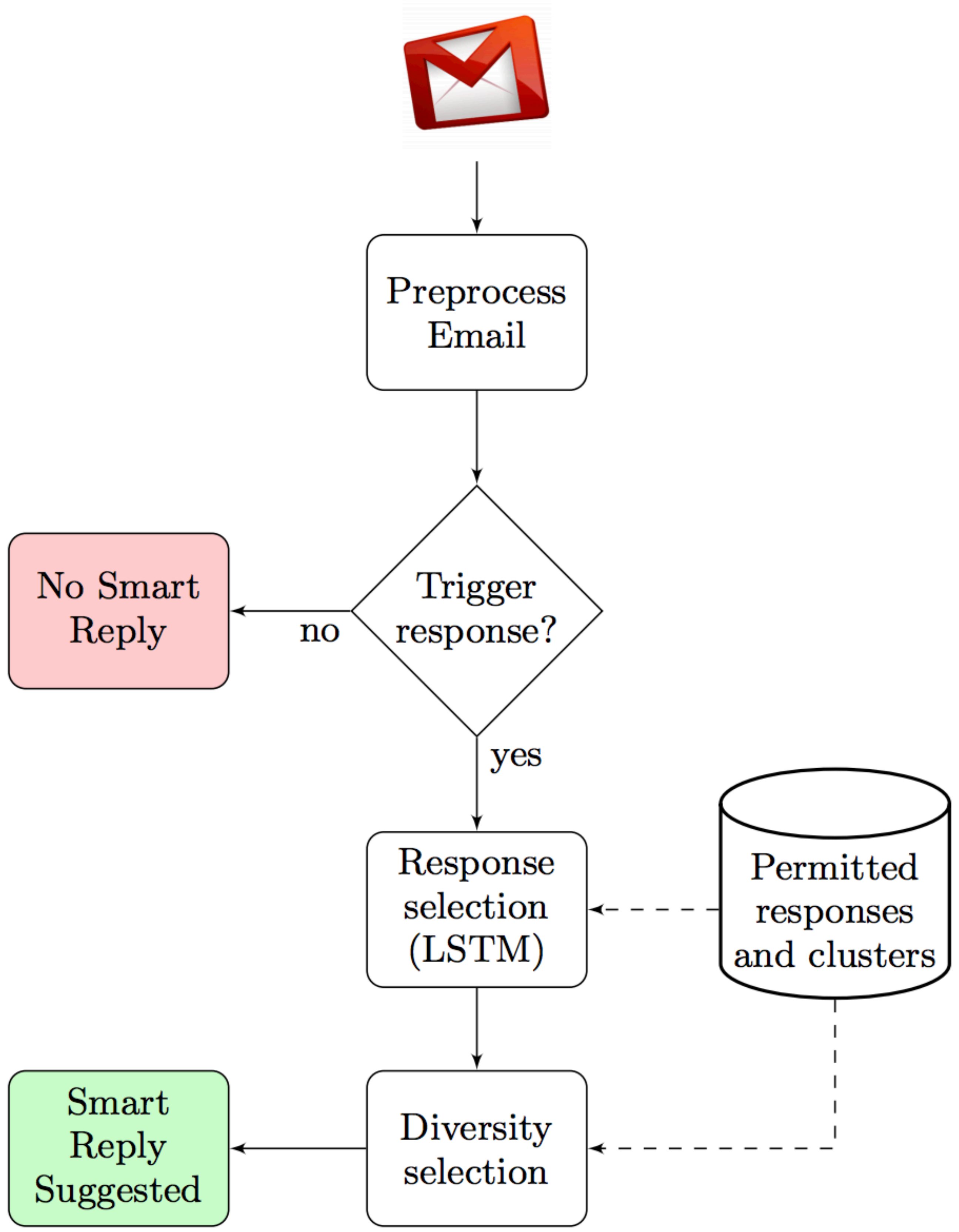
Wen et al, 2018 (to be appear)

Case Study: Google Smart Reply

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

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

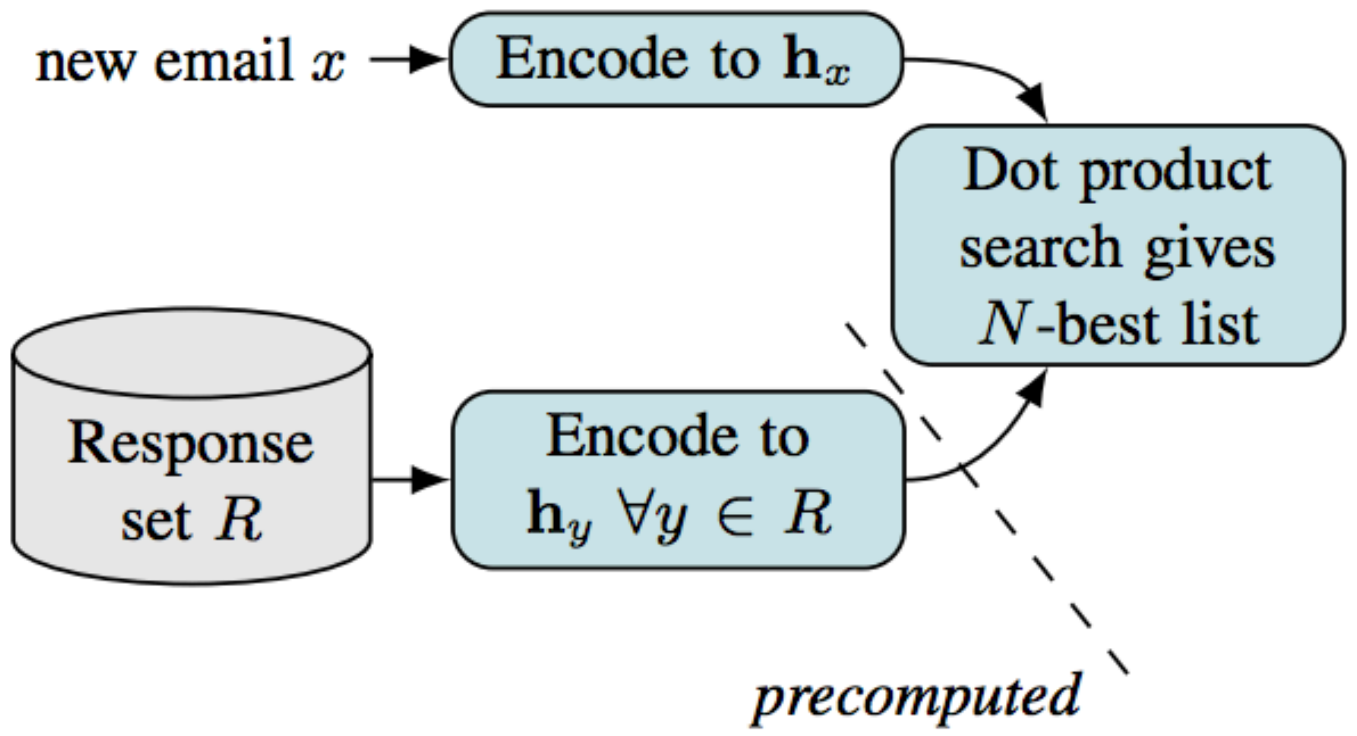
- Following the Seq2Seq paradigm.
- Inference is done by:
  - Organizing responses as a trie
  - Beam search



Kannan et al, 2016

Case Study: Google Smart Reply

- Encode response set (precomputed)
- Dot product (fast) between email and response vectors.
- Hierarchical quantization for search.


new email x
Encode to $\mathbf{h}_x$
Dot product search gives N-best list
Response set R
Encode to $\mathbf{h}_y \ \forall y \in R$
precomputed

	Conversion Rate (relative to Seq2Seq)	Speed (relative to Seq2Seq)
Seq2Seq + beam search	100%	1x
Dot Product + hierarchical search	104%	**100x**

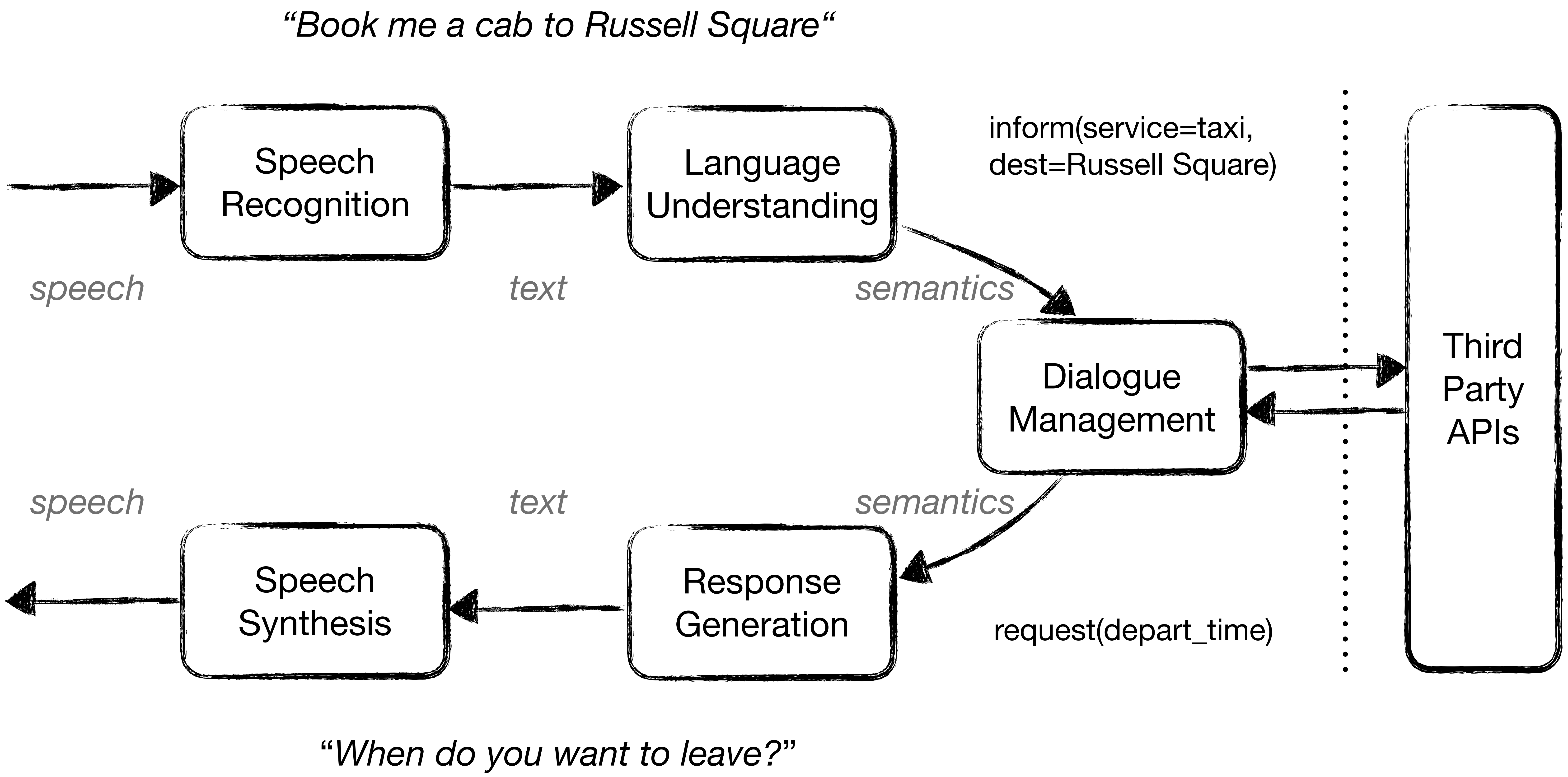
Henderson et al, 2017

Short Summary

- Generative Sequence-to-Sequence model is extremely hard to get it to work.
    - It struggled in generating diverse responses - but other methods have been proposed to address this problem.
    - Generating the intended meaning while learning the semantic representation on the fly is still impossible.
- Runtime-wise, generative model is much slower than simple dot product retrieval-based model.
- Dot product model has been successfully deployed and shows a better result than generative models.

Learning-based
Goal-oriented Dialogue

Recap - Conversational AI Architecture



“Book me a cab to Russell Square“

Speech Recognition

Language Understanding

inform(service=taxi,
dest=Russell Square)

speech

text

semantics

Dialogue Management

Third Party APIs

speech

text

semantics

Speech Synthesis

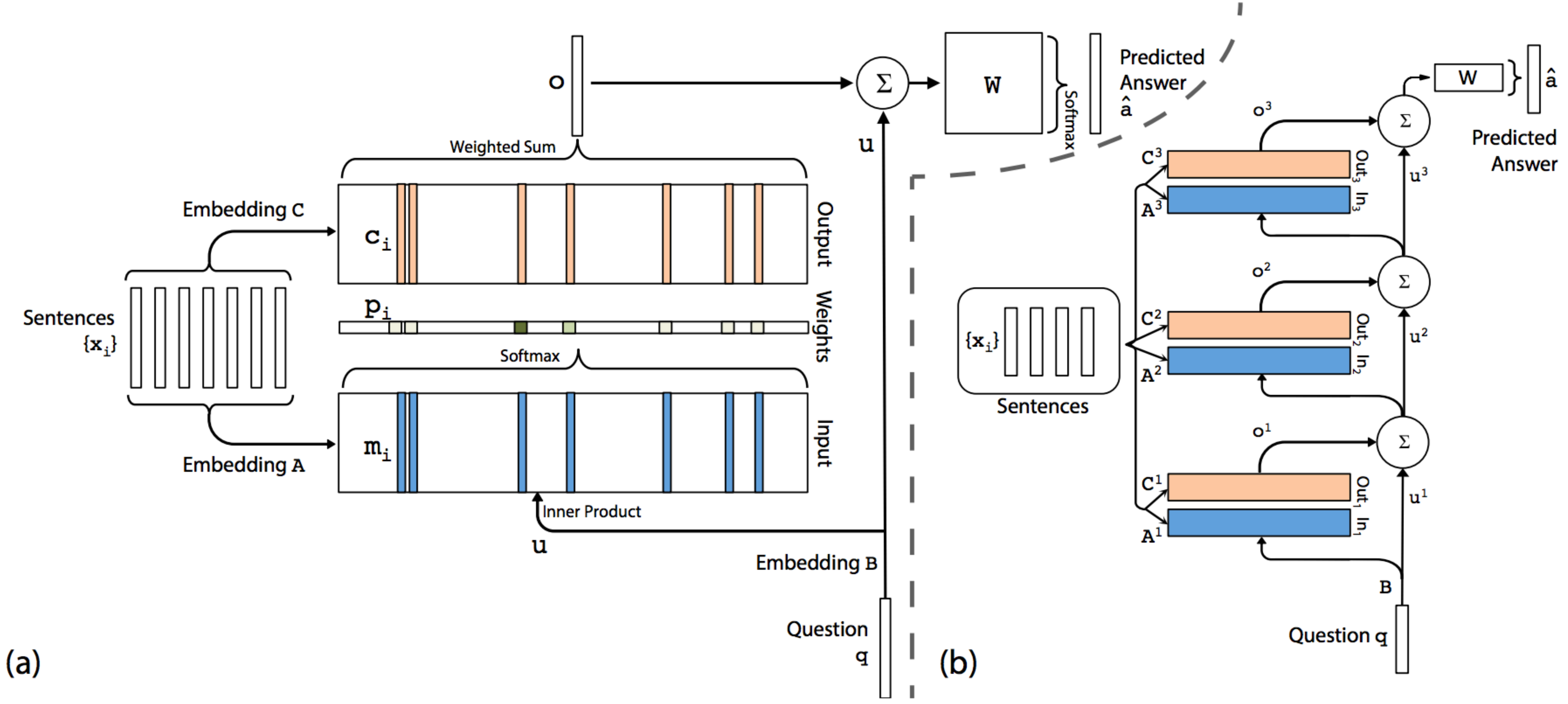
Response Generation

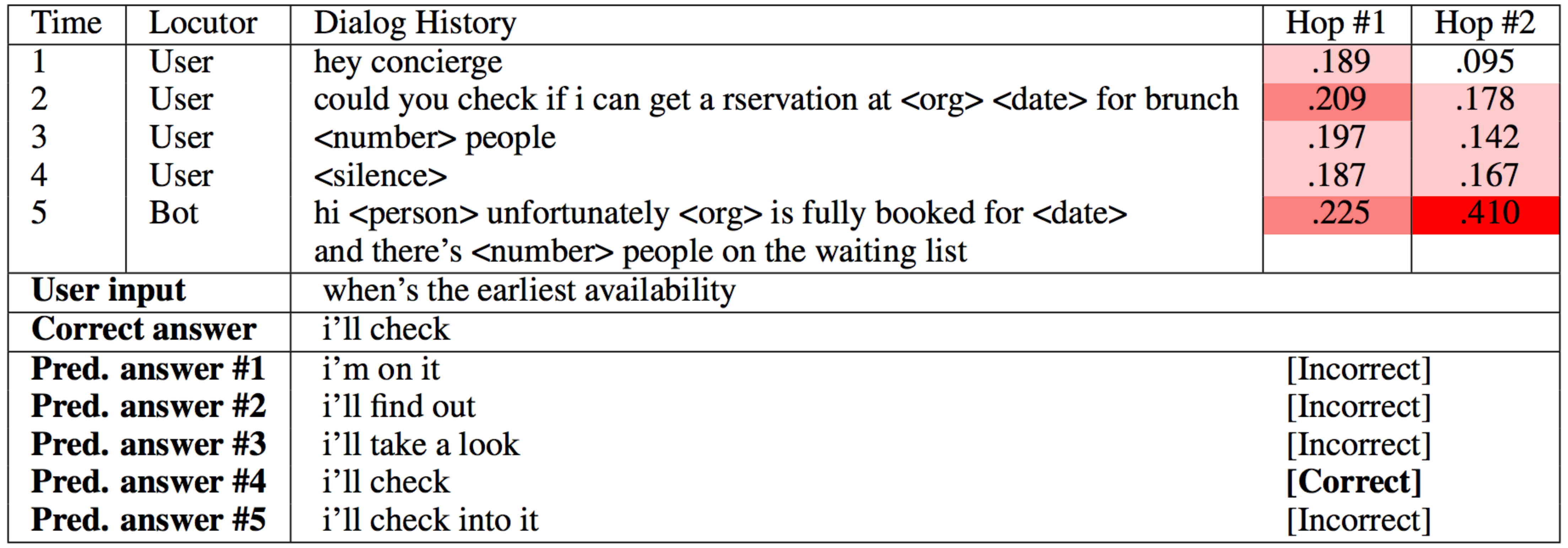
request(depart_time)

“When do you want to leave?”

Memory Network for Dialogue Modeling

Bordes et al, ICLR 2017

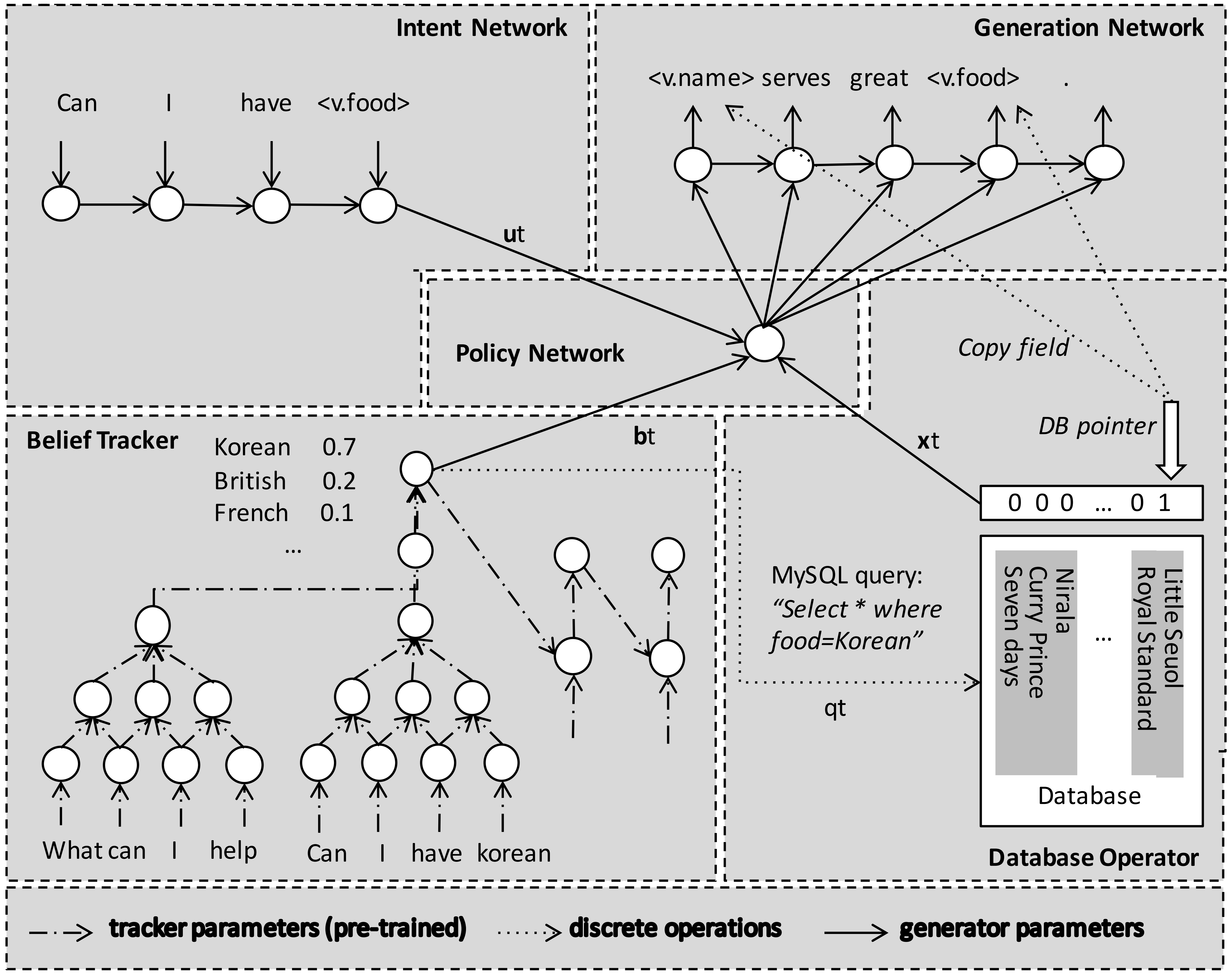




Time	Locutor	Dialog History	Hop #1	Hop #2
1	User	hey concierge	.189	.095
2	User	could you check if i can get a rservation at <org> <date> for brunch	.209	.178
3	User	<number> people	.197	.142
4	User	<silence>	.187	.167
5	Bot	hi <person> unfortunately <org> is fully booked for <date> and there’s <number> people on the waiting list	.225	.410
User input		when’s the earliest availability		
Correct answer		i’ll check		
Pred. answer #1		i’m on 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]	

- 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.

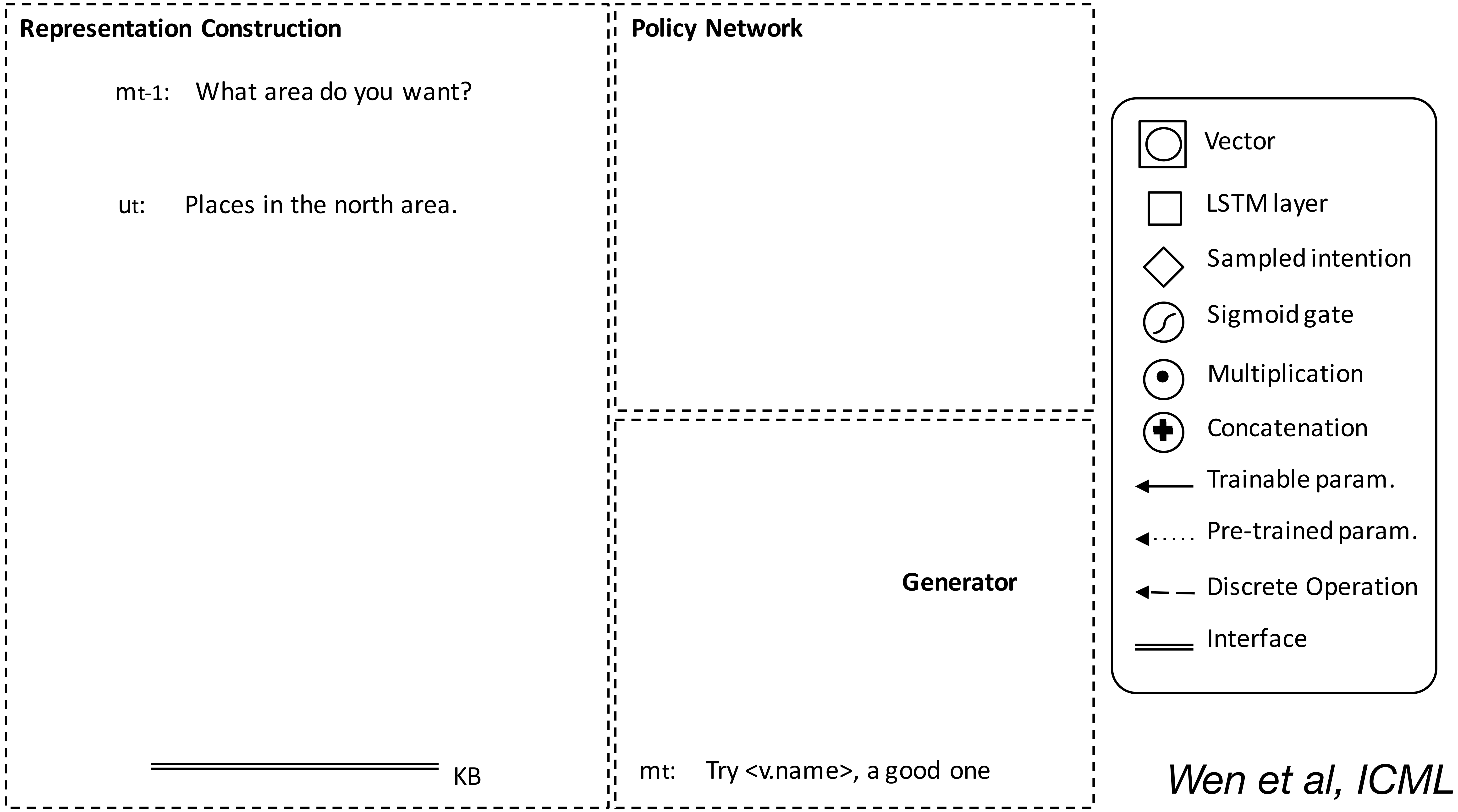
Neural Dialogue Model



Wen et al, EACL 2017

- Explicitly use the structure of dialogue systems.
- Adopt slot-value pairs and domain-specific ontology in the development.
- A discrete access to database based on the parsed intent.
- Apply neural networks to predict intent and generate responses.
- Task success rate is around 92% when trained on only 400 dialogues.

Latent Intention Dialogue Model



Representation Construction

m_{t-1}: What area do you want?

u_t: Places in the north area.

KB

Policy Network

Generator

m_t: Try <v.name>, a good one

Vector

LSTM layer

Sampled intention

Sigmoid gate

Multiplication

Concatenation

Trainable param.

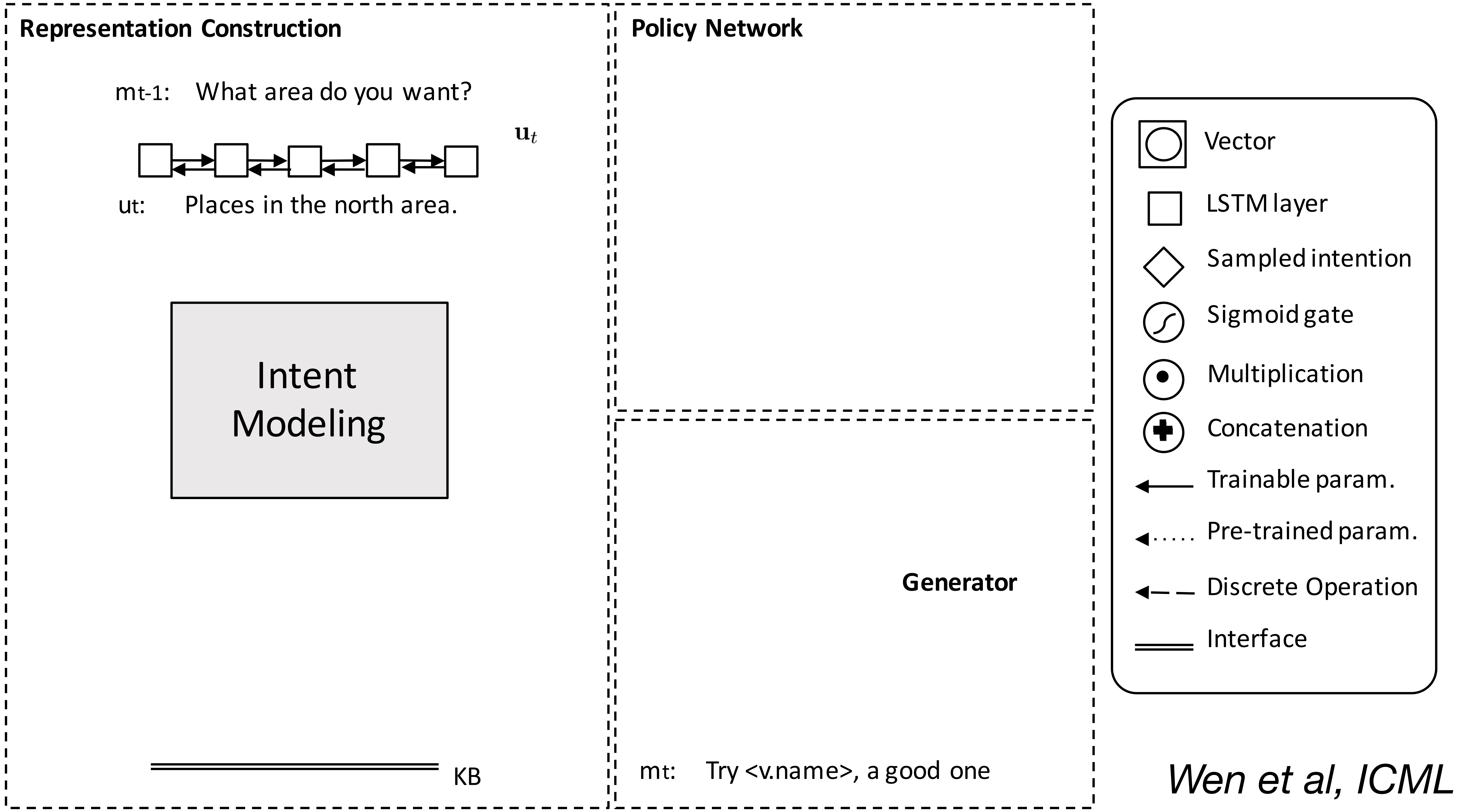
Pre-trained param.

Discrete Operation

Interface

Wen et al, ICML 2017.

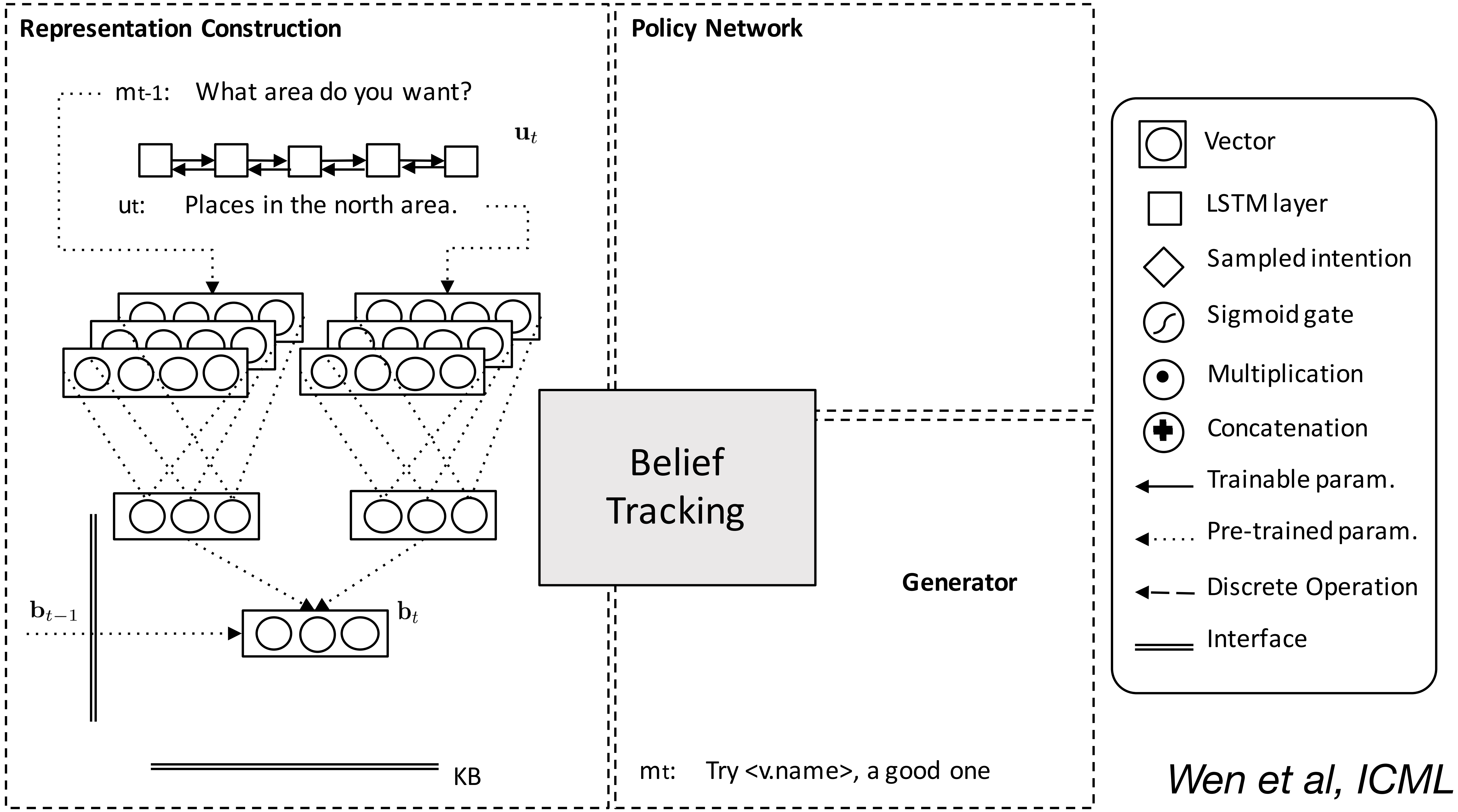
Latent Intention Dialogue Model



Wen et al, ICML 2017.

$$\mathbf{u}_t = \text{biLSTM}(u_t)$$

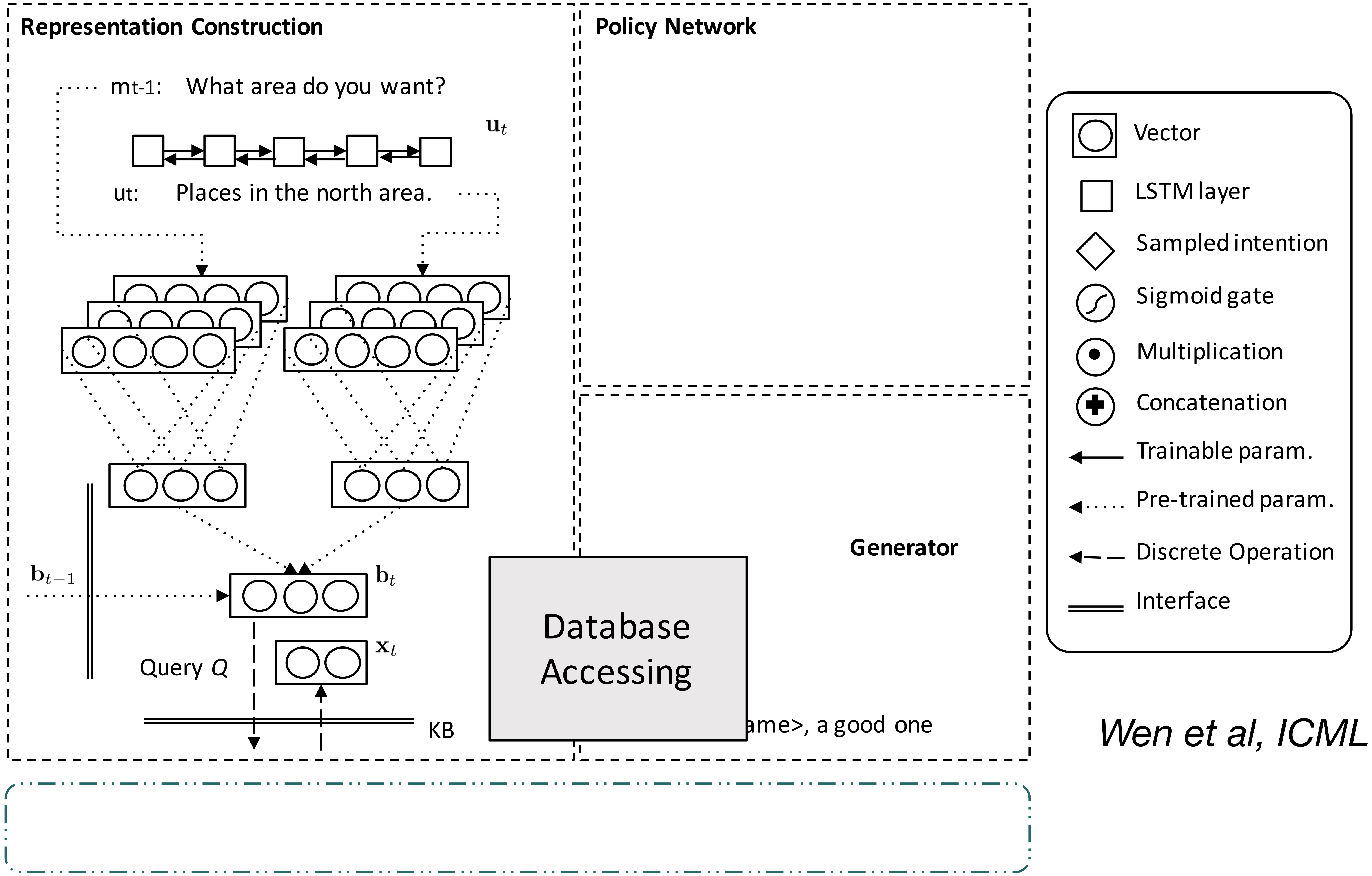
Latent Intention Dialogue Model



Wen et al, ICML 2017.

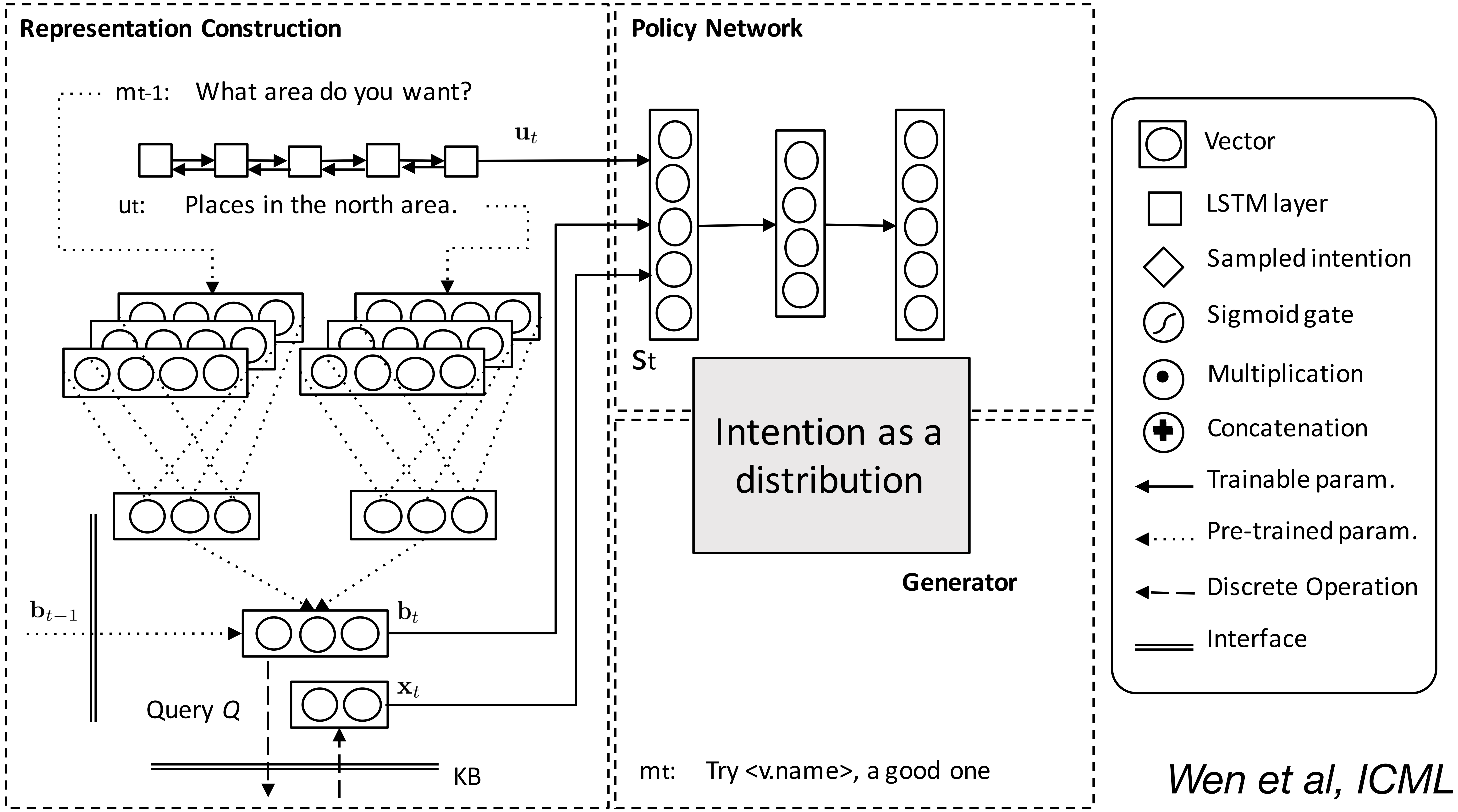
$$\mathbf{b}_t = \text{RNN-CNN}(u_t, m_{t-1}, \mathbf{b}_{t-1})$$

Latent Intention Dialogue Model



Wen et al, ICML 2017.

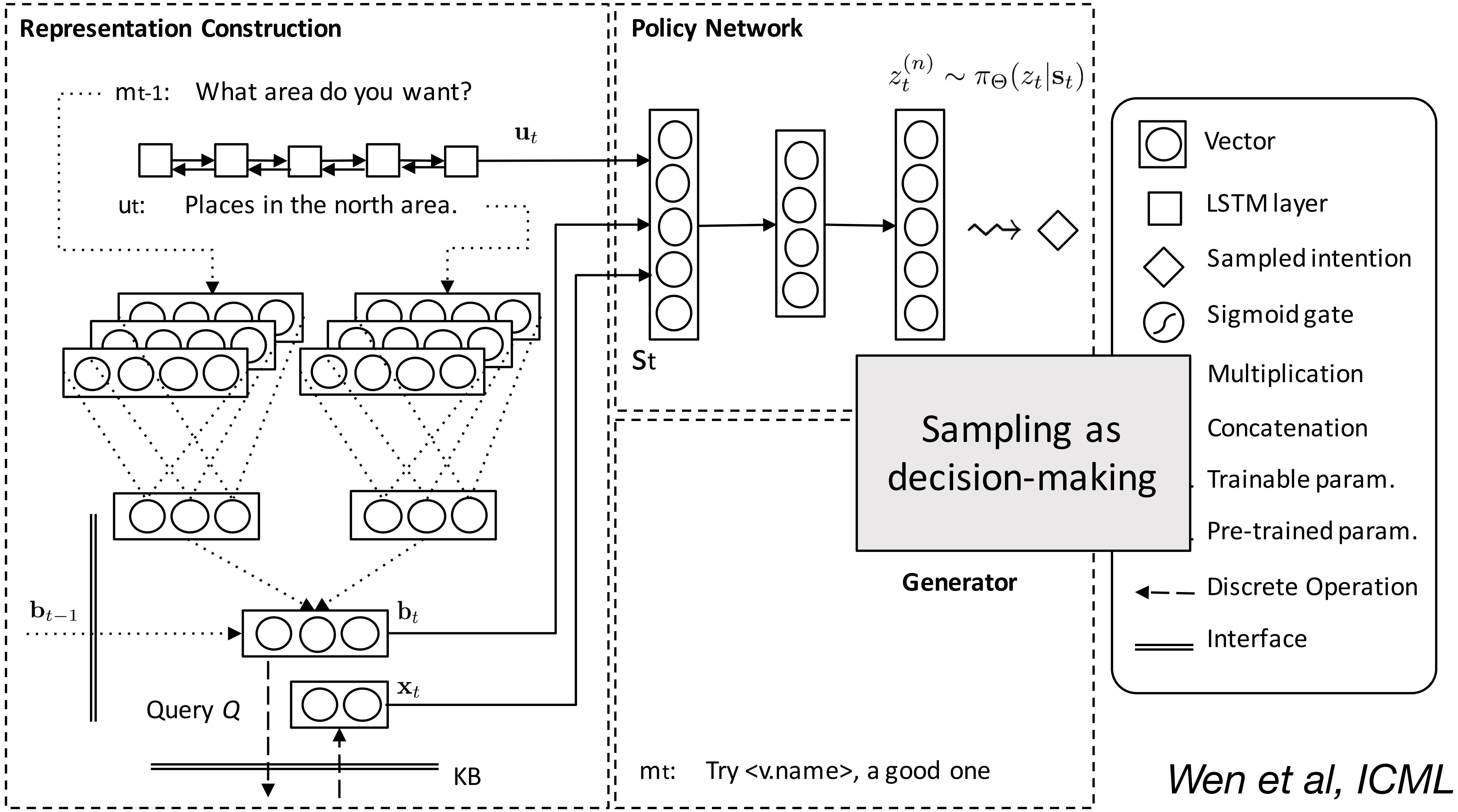
Latent Intention Dialogue Model



Wen et al, ICML 2017.

$$\mathbf{s}_t = \mathbf{u}_t \oplus \mathbf{b}_t \oplus \mathbf{x}_t \qquad \pi_\theta(z_t|\mathbf{s}_t) = \mathrm{softmax}(\mathbf{W}_2^\intercal \cdot \tanh(\mathbf{W}_1^\intercal \mathbf{s}_t + \mathbf{b}_1) + \mathbf{b}_2)$$

Latent Intention Dialogue Model

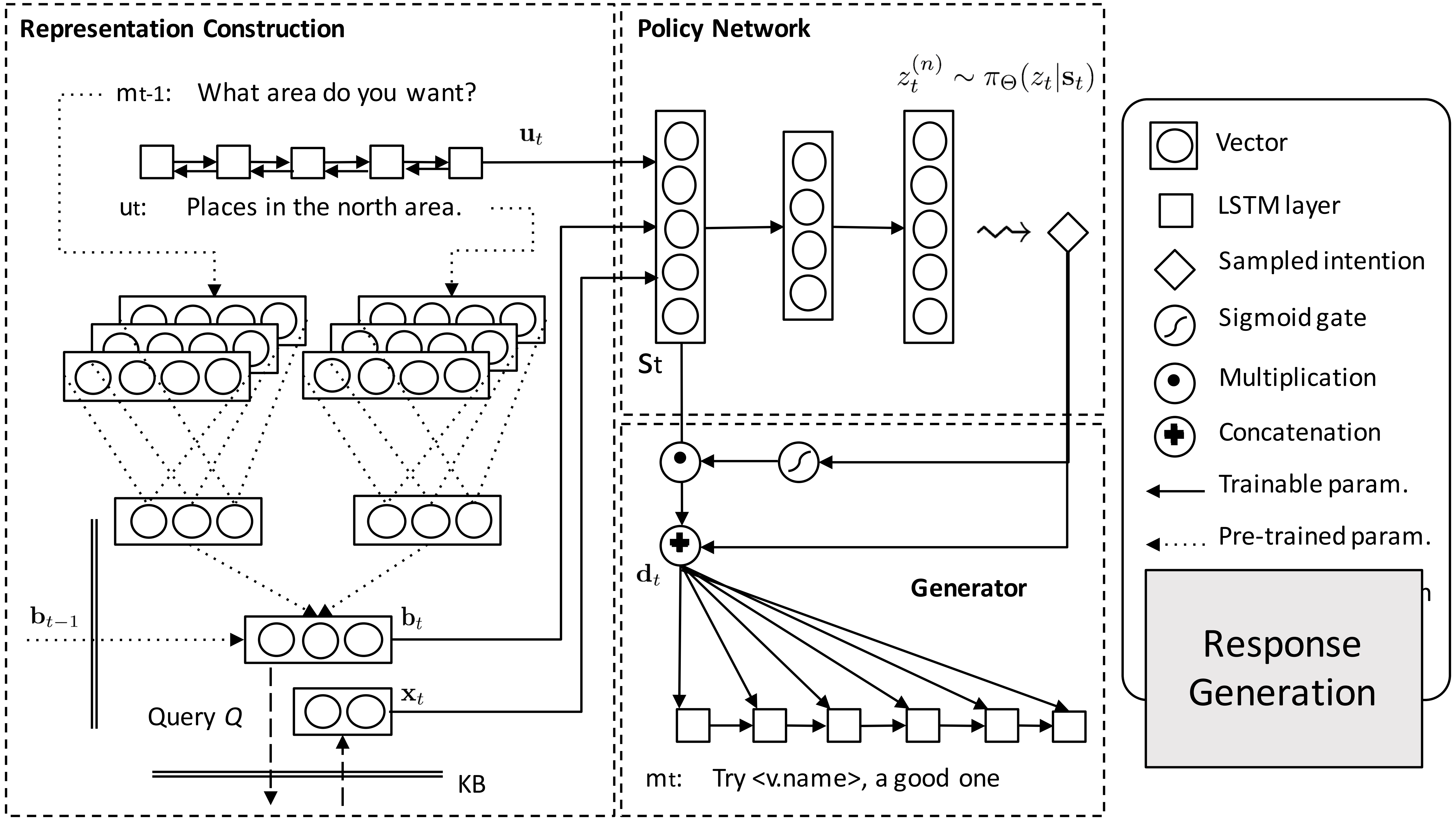


Sampling as decision-making

Wen et al, ICML 2017.

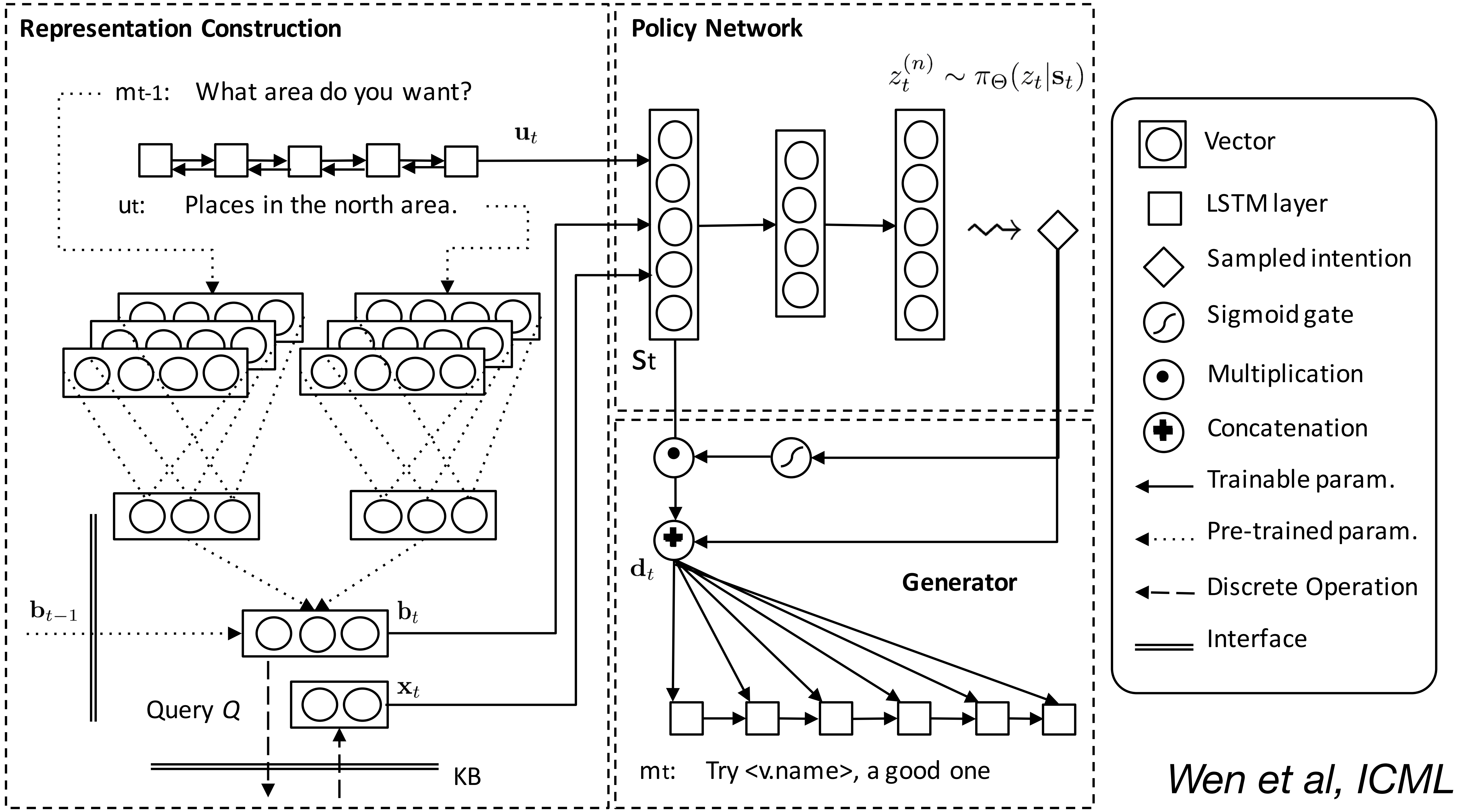
$$z_t^{(n)} \sim \pi_\theta(z_t|\mathbf{s}_t)$$

Latent Intention Dialogue Model



$$\mathbf{d}_t = \mathbf{W}_4^{\intercal}\mathbf{z}_t \oplus \left[\text{sigmoid}(\mathbf{W}_3^{\intercal}\mathbf{z}_t + \mathbf{b}_3) \cdot \mathbf{W}_5^{\intercal}\mathbf{s}_t\right] \qquad p_\theta(m_t|\mathbf{s}_t, z_t) = \prod_j p(w_{j+1}^t | w_j^t, \mathbf{h}_{j-1}^t, \mathbf{d}_t)$$

Latent Intention Dialogue Model



Wen et al, ICML 2017.

$$p_\theta(m_t|\mathbf{s}_t) = \sum_{z_t} p_\theta(m_t|z_t,\mathbf{s}_t)\pi_\theta(z_t|\mathbf{s}_t)$$

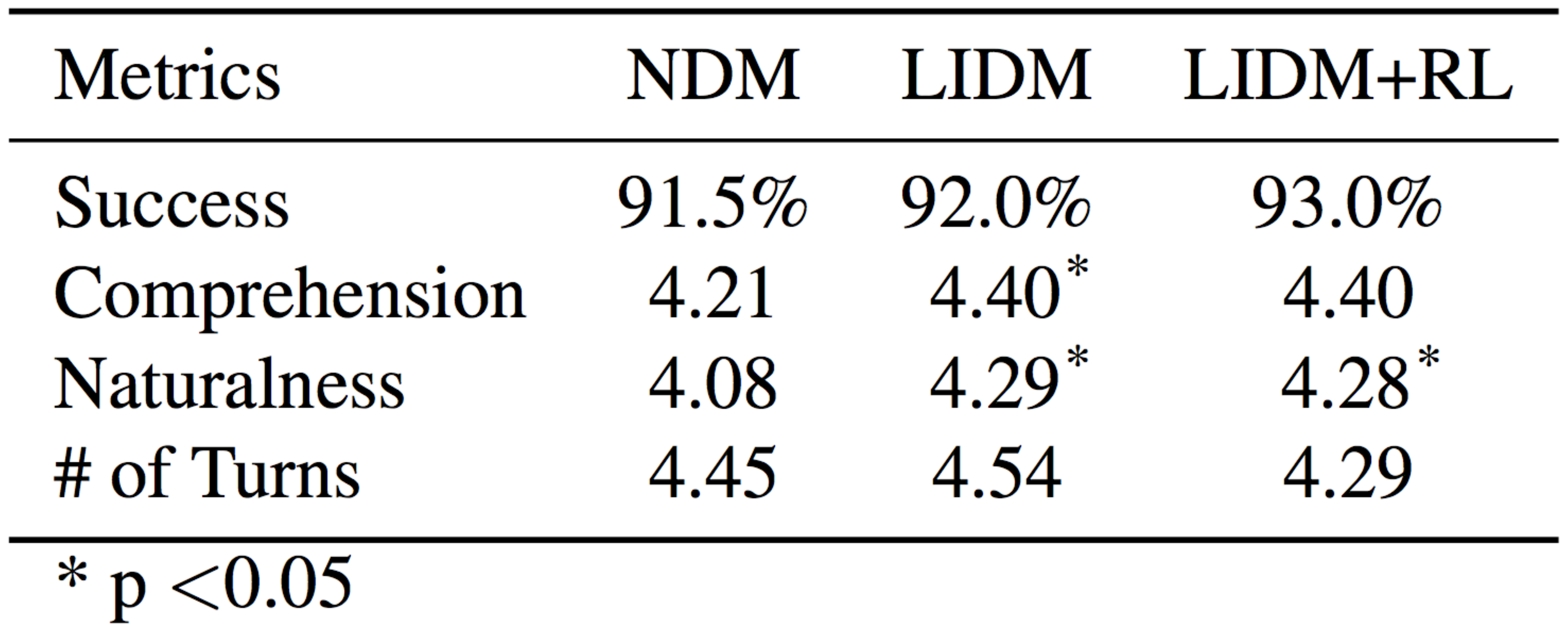
Learning of LIDM

- Approximate the objective with a variational lowerbound:

$$\begin{aligned}\mathcal{L}(\theta,\phi) &= \mathbb{E}_{q_\phi(z_t)}[\log p_\theta(m_t|z_t,\mathbf{s}_t)] - \lambda D_{KL}(q_\phi(z_t)||\pi_\theta(z_t|\mathbf{s}_t)) \\ &\leq \log \sum_{z_t} p_\theta(m_t|z_t,\mathbf{s}_t)\pi_\theta(z_t|\mathbf{s}_t) \\ &= \log p_\theta(m_t|\mathbf{s}_t)\end{aligned}$$

- Follow the **Variational Inference** method.
- *Latent Intention Dialogue Models, Wen et al, ICML 2017*

Human Evaluation

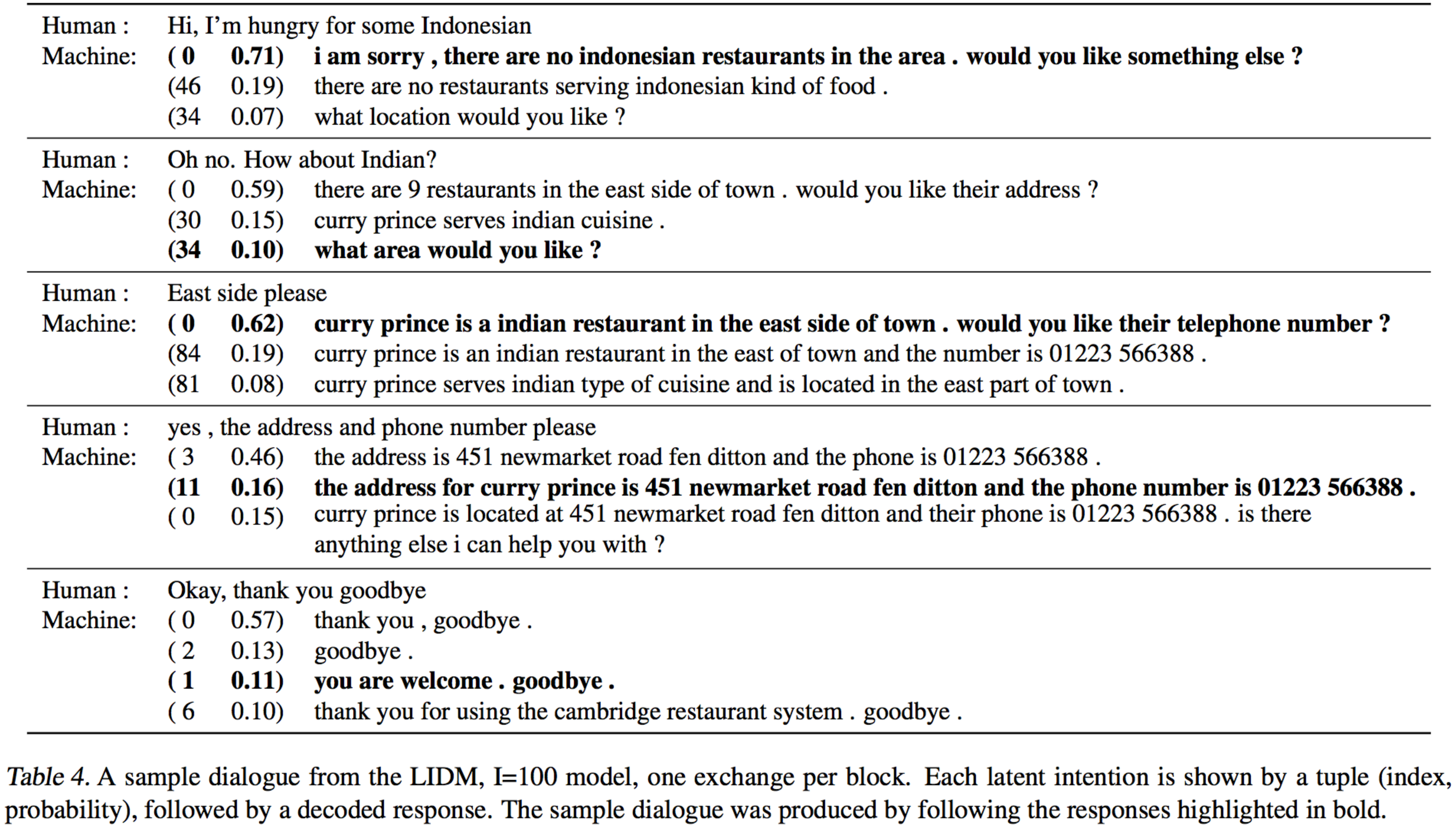


Metrics	NDM	LIDM	LIDM+RL
Success	91.5%	92.0%	93.0%
Comprehension	4.21	4.40*	4.40
Naturalness	4.08	4.29*	4.28*
# of Turns	4.45	4.54	4.29

* $p < 0.05$

- A high dialogue success rate: >91%.
- Comprehension and naturalness are both rated high >4/5

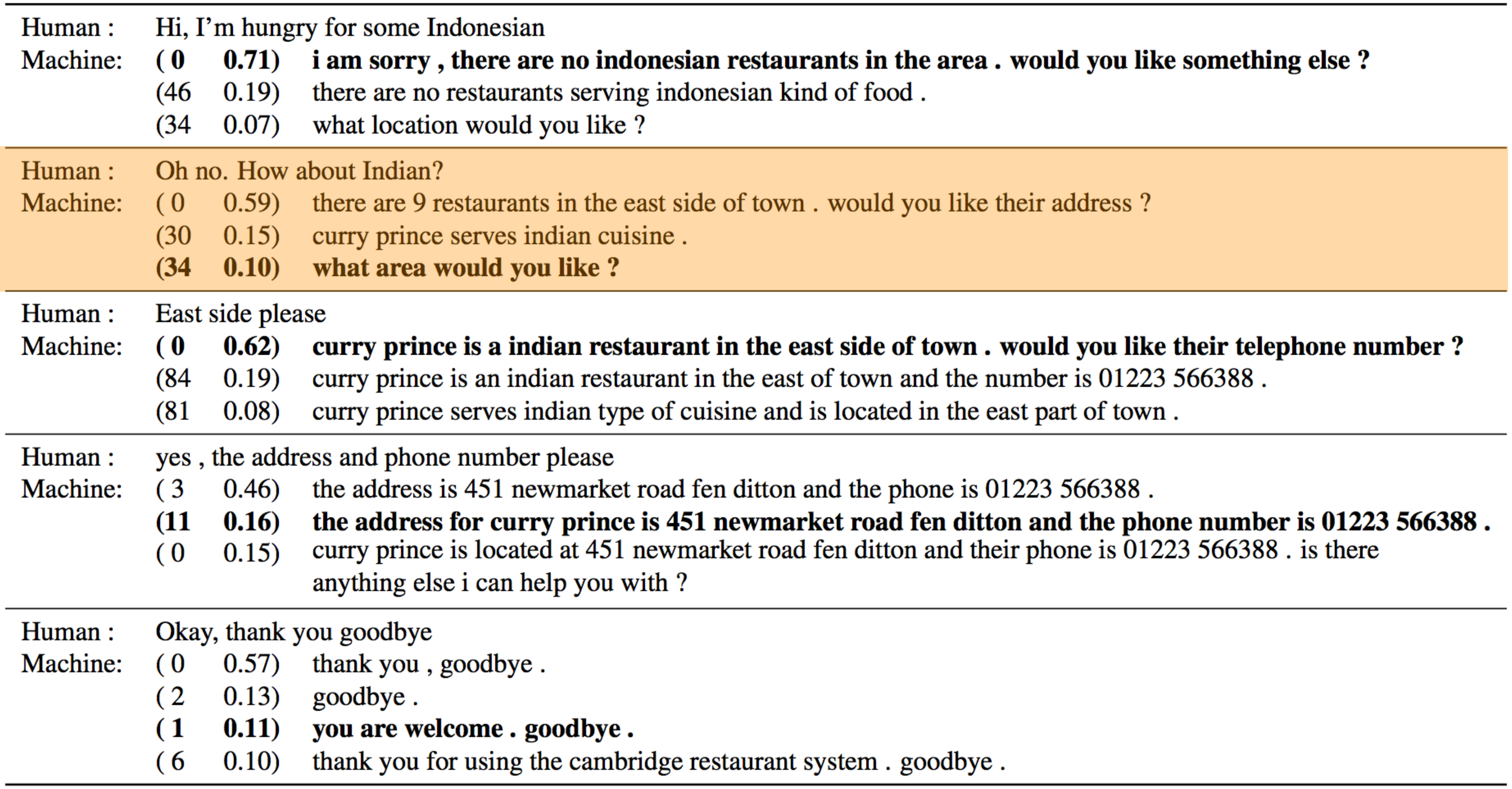
Sample Conversation



Human :	Hi, I'm hungry for some Indonesian		
Machine:	**(0**	**0.71)**	**i am sorry , there are no indonesian restaurants in the area . would you like something else ?**
	(46	0.19)	there are no restaurants serving indonesian kind of food .
	(34	0.07)	what location would you like ?
Human :	Oh no. How about Indian?		
Machine:	(0	0.59)	there are 9 restaurants in the east side of town . would you like their address ?
	(30	0.15)	curry prince serves indian cuisine .
	(34	**0.10)**	**what area would you like ?**
Human :	East side please		
Machine:	**(0**	**0.62)**	**curry prince is a indian restaurant in the east side of town . would you like their telephone number ?**
	(84	0.19)	curry prince is an indian restaurant in the east of town and the number is 01223 566388 .
	(81	0.08)	curry prince serves indian type of cuisine and is located in the east part of town .
Human :	yes , the address and phone number please		
Machine:	(3	0.46)	the address is 451 newmarket road fen ditton and the phone is 01223 566388 .
	(11	**0.16)**	**the address for curry prince is 451 newmarket road fen ditton and the phone number is 01223 566388 .**
	(0	0.15)	curry prince is located at 451 newmarket road fen ditton and their phone is 01223 566388 . is there anything else i can help you with ?
Human :	Okay, thank you goodbye		
Machine:	(0	0.57)	thank you , goodbye .
	(2	0.13)	goodbye .
	(1	**0.11)**	**you are welcome . goodbye .**
	(6	0.10)	thank you for using the cambridge restaurant system . goodbye .

Table 4. A sample dialogue from the LIDM, I=100 model, one exchange per block. Each latent intention is shown by a tuple (index, probability), followed by a decoded response. The sample dialogue was produced by following the responses highlighted in bold.

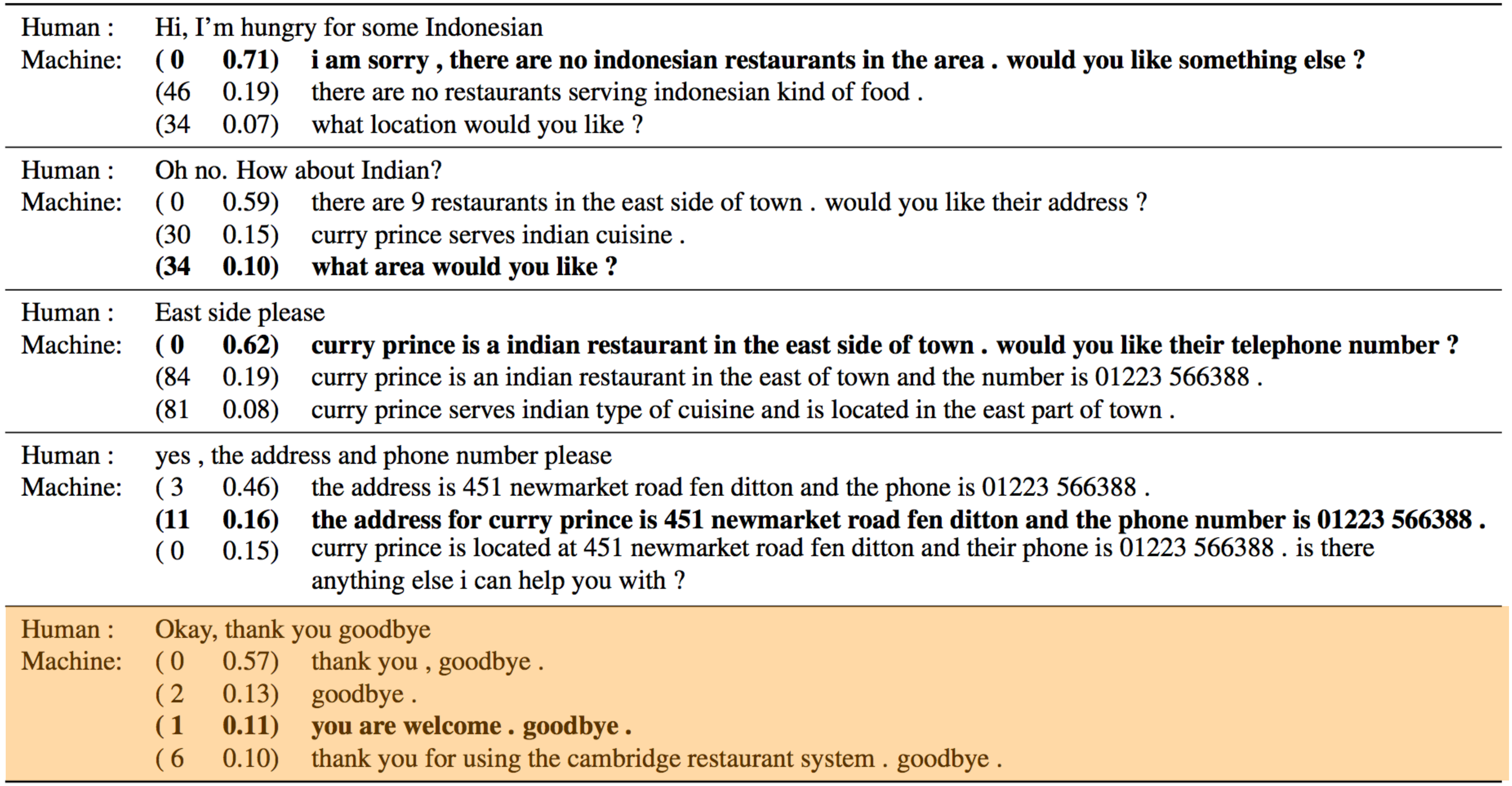
Sample Conversation



Human :	Hi, I'm hungry for some Indonesian		
Machine:	**(0**	**0.71)**	**i am sorry , there are no indonesian restaurants in the area . would you like something else ?**
	(46	0.19)	there are no restaurants serving indonesian kind of food .
	(34	0.07)	what location would you like ?
Human :	Oh no. How about Indian?		
Machine:	(0	0.59)	there are 9 restaurants in the east side of town . would you like their address ?
	(30	0.15)	curry prince serves indian cuisine .
	(34	**0.10)**	**what area would you like ?**
Human :	East side please		
Machine:	**(0**	**0.62)**	**curry prince is a indian restaurant in the east side of town . would you like their telephone number ?**
	(84	0.19)	curry prince is an indian restaurant in the east of town and the number is 01223 566388 .
	(81	0.08)	curry prince serves indian type of cuisine and is located in the east part of town .
Human :	yes , the address and phone number please		
Machine:	(3	0.46)	the address is 451 newmarket road fen ditton and the phone is 01223 566388 .
	(11	**0.16)**	**the address for curry prince is 451 newmarket road fen ditton and the phone number is 01223 566388 .**
	(0	0.15)	curry prince is located at 451 newmarket road fen ditton and their phone is 01223 566388 . is there anything else i can help you with ?
Human :	Okay, thank you goodbye		
Machine:	(0	0.57)	thank you , goodbye .
	(2	0.13)	goodbye .
	(1	**0.11)**	**you are welcome . goodbye .**
	(6	0.10)	thank you for using the cambridge restaurant system . goodbye .

Table 4. A sample dialogue from the LIDM, I=100 model, one exchange per block. Each latent intention is shown by a tuple (index, probability), followed by a decoded response. The sample dialogue was produced by following the responses highlighted in bold.

Sample Conversation

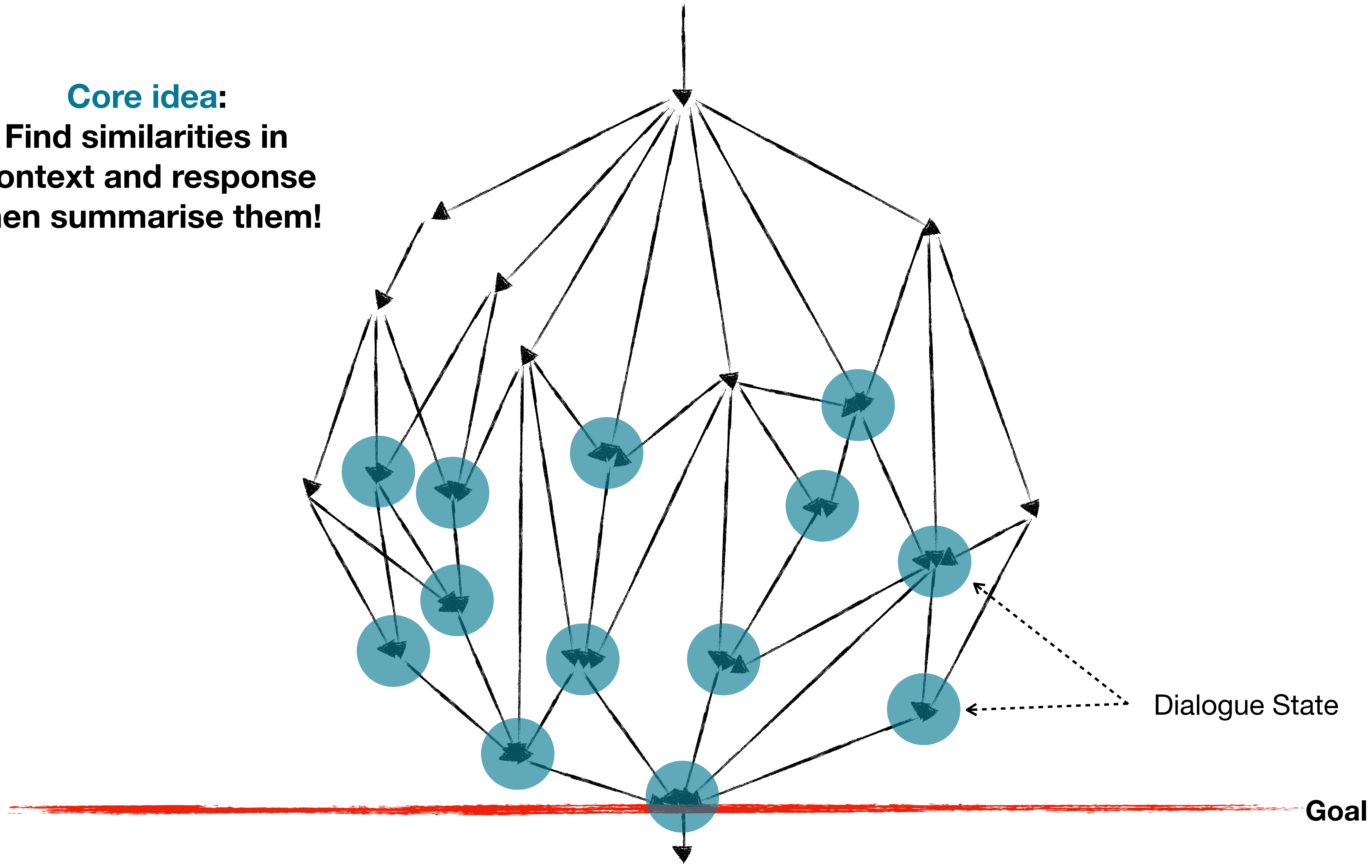


Human :	Hi, I'm hungry for some Indonesian		
Machine:	**(0**	**0.71)**	**i am sorry , there are no indonesian restaurants in the area . would you like something else ?**
	(46	0.19)	there are no restaurants serving indonesian kind of food .
	(34	0.07)	what location would you like ?
Human :	Oh no. How about Indian?		
Machine:	(0	0.59)	there are 9 restaurants in the east side of town . would you like their address ?
	(30	0.15)	curry prince serves indian cuisine .
	(34	**0.10)**	**what area would you like ?**
Human :	East side please		
Machine:	**(0**	**0.62)**	**curry prince is a indian restaurant in the east side of town . would you like their telephone number ?**
	(84	0.19)	curry prince is an indian restaurant in the east of town and the number is 01223 566388 .
	(81	0.08)	curry prince serves indian type of cuisine and is located in the east part of town .
Human :	yes , the address and phone number please		
Machine:	(3	0.46)	the address is 451 newmarket road fen ditton and the phone is 01223 566388 .
	(11	**0.16)**	**the address for curry prince is 451 newmarket road fen ditton and the phone number is 01223 566388 .**
	(0	0.15)	curry prince is located at 451 newmarket road fen ditton and their phone is 01223 566388 . is there anything else i can help you with ?
Human :	Okay, thank you goodbye		
Machine:	(0	0.57)	thank you , goodbye .
	(2	0.13)	goodbye .
	(1	**0.11)**	**you are welcome . goodbye .**
	(6	0.10)	thank you for using the cambridge restaurant system . goodbye .

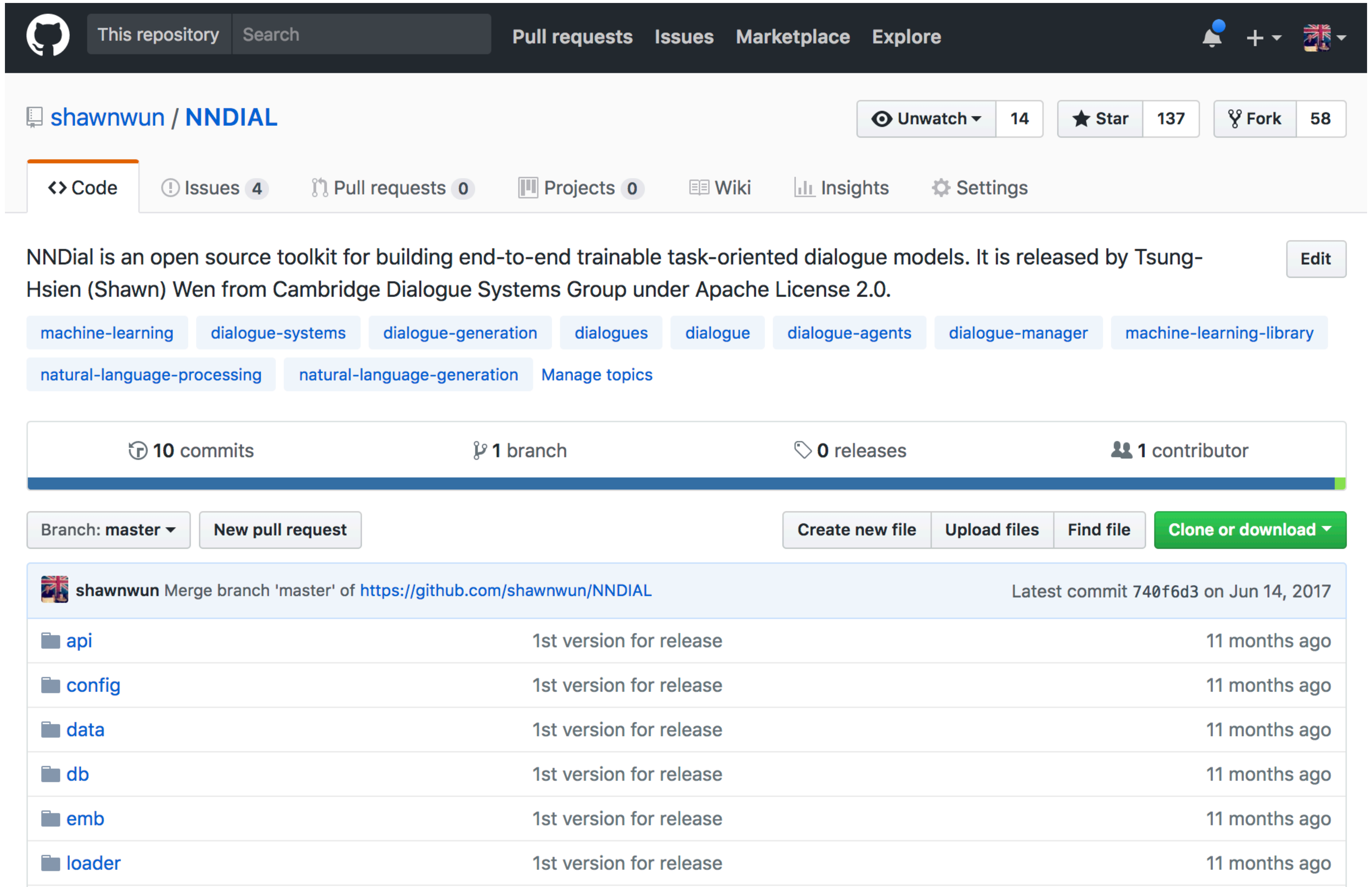
Table 4. A sample dialogue from the LIDM, I=100 model, one exchange per block. Each latent intention is shown by a tuple (index, probability), followed by a decoded response. The sample dialogue was produced by following the responses highlighted in bold.

The PolyAI Approach

Core idea:
Find similarities in context and response then summarise them!



NNDIAL Toolkit



shawnwun / **NNDIAL**

NNDial is an open source toolkit for building end-to-end trainable task-oriented dialogue models. It is released by Tsung-Hsien (Shawn) Wen from Cambridge Dialogue Systems Group under Apache License 2.0.

shawnwun Merge branch 'master' of https://github.com/shawnwun/NNDIAL Latest commit 740f6d3 on Jun 14, 2017

api	1st version for release	11 months ago
config	1st version for release	11 months ago
data	1st version for release	11 months ago
db	1st version for release	11 months ago
emb	1st version for release	11 months ago
loader	1st version for release	11 months ago

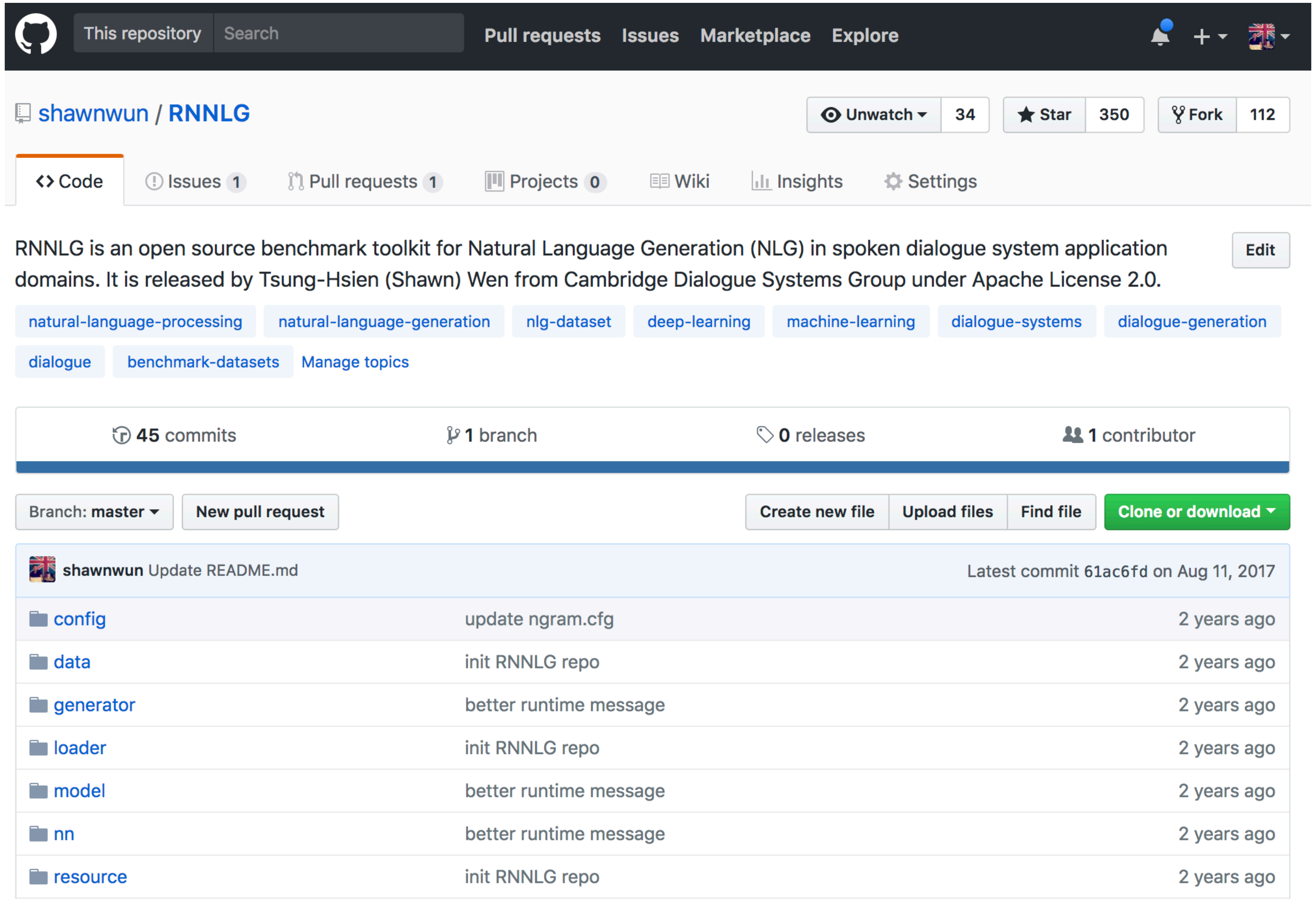
Link:
https://github.com/shawnwun/NNDIAL

A toolkit for training neural network based dialogue models.

Include both NDM, Attentive NDM, and LIDM implementations.

Include a CamRest676 dialogue dataset.

RNNLG Toolkit



RNNLG is an open source benchmark toolkit for Natural Language Generation (NLG) in spoken dialogue system application domains. It is released by Tsung-Hsien (Shawn) Wen from Cambridge Dialogue Systems Group under Apache License 2.0.

shawnwun Update README.md — Latest commit `61ac6fd` on Aug 11, 2017

config	update ngram.cfg	2 years ago
data	init RNNLG repo	2 years ago
generator	better runtime message	2 years ago
loader	init RNNLG repo	2 years ago
model	better runtime message	2 years ago
nn	better runtime message	2 years ago
resource	init RNNLG repo	2 years ago

Link:
https://github.com/shawnwun/RNNLG

A toolkit for rendering text from a dialogue act representation.

Include several RNN language generator implementation like Attention-based Encoder Decoder, SC-LSTM, H-LSTM … etc.

Include datasets from four different application domains.

Short Summary

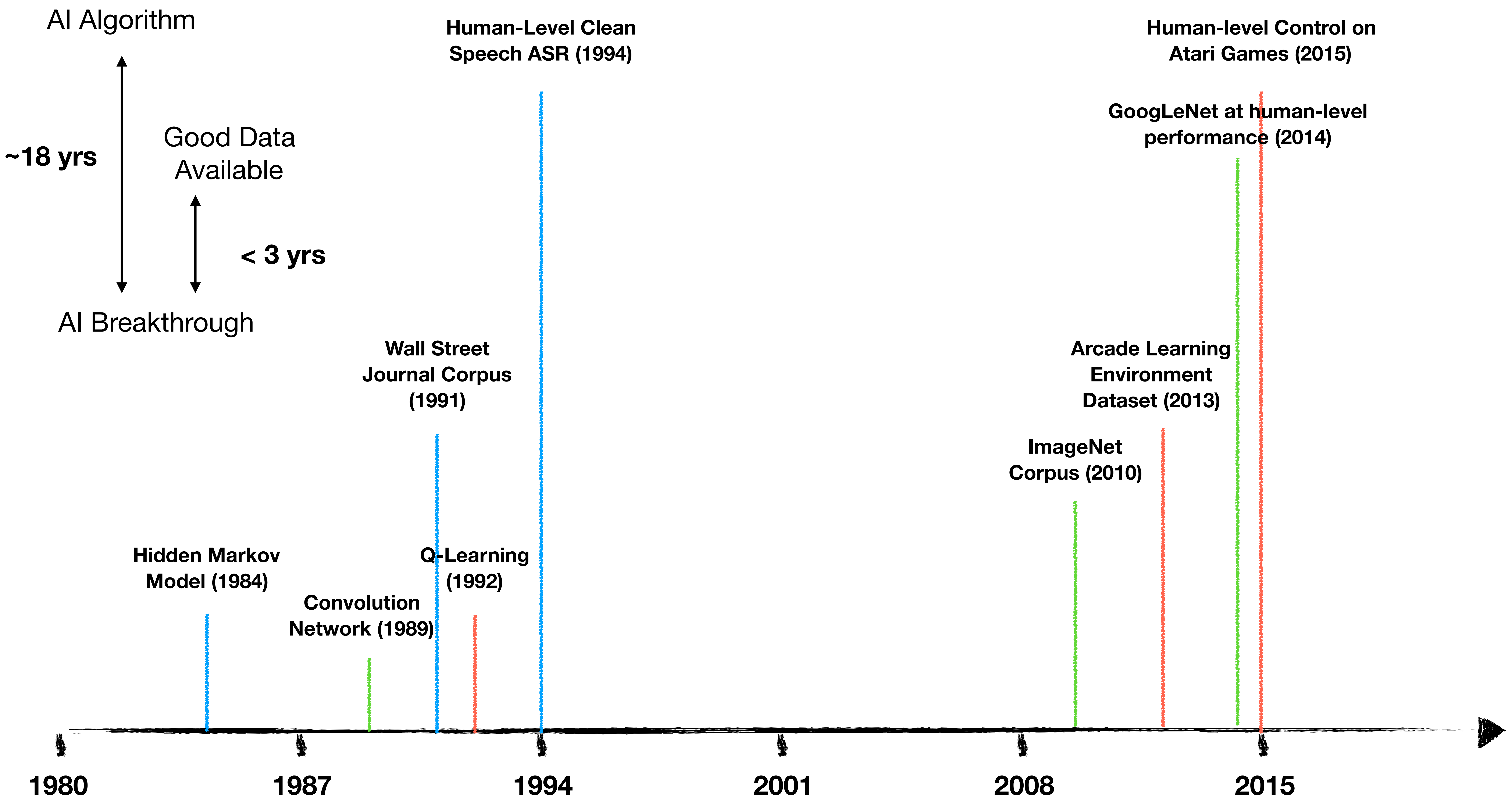
- Conversational AI is a collective effort of multiple parties and requires a mix of multiple methods.
- Machine learning is
    - a perfect tool for understanding quries.
    - an OK method for response selection/generation.
    - but terrible for dialogue decision-making and database integration (the training data limitation).
- Practical but open-minded in choosing approaches is the way to build robust and truly useful conversational AIs of the future.

Data! Data! Data!

The Importance of Data

- Data is the “*core*” of ML algorithms.
  - One of the reasons why Deep Learning took over.
- Data should be the “*core*” of an ML company too.
  - You should own your data.
  - You should have a way to generate your data.
  - You should have a way to label your data.

The Importance of Data

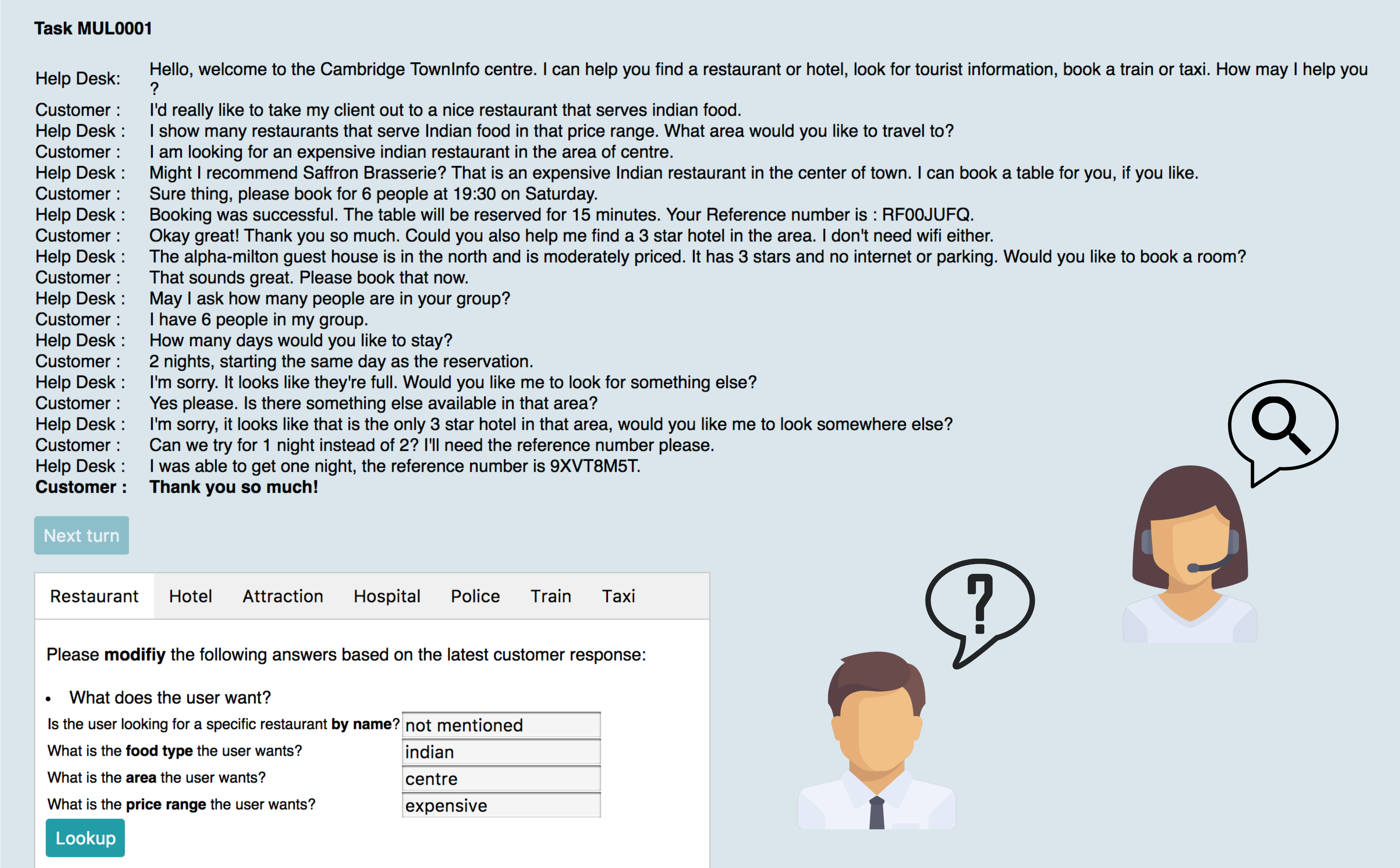


Reference: http://www.spacemachine.net/views/2016/3/datasets-over-algorithms

The Importance of Data

- Conversational data is hard to get.
- Reddit has a large volume of conversational data.
  - For chit-chat dialogue development.
- But how do we get domain specific datasets?
  - A pizza delivery agent?
  - A taxi booking bot?
- COLLECT YOUR OWN DATA!

Example: WOZ Data Collection



Task MUL0001

Help Desk: Hello, welcome to the Cambridge TownInfo centre. I can help you find a restaurant or hotel, look for tourist information, book a train or taxi. How may I help you ?
Customer : I'd really like to take my client out to a nice restaurant that serves indian food.
Help Desk : I show many restaurants that serve Indian food in that price range. What area would you like to travel to?
Customer : I am looking for an expensive indian restaurant in the area of centre.
Help Desk : Might I recommend Saffron Brasserie? That is an expensive Indian restaurant in the center of town. I can book a table for you, if you like.
Customer : Sure thing, please book for 6 people at 19:30 on Saturday.
Help Desk : Booking was successful. The table will be reserved for 15 minutes. Your Reference number is : RF00JUFQ.
Customer : Okay great! Thank you so much. Could you also help me find a 3 star hotel in the area. I don't need wifi either.
Help Desk : The alpha-milton guest house is in the north and is moderately priced. It has 3 stars and no internet or parking. Would you like to book a room?
Customer : That sounds great. Please book that now.
Help Desk : May I ask how many people are in your group?
Customer : I have 6 people in my group.
Help Desk : How many days would you like to stay?
Customer : 2 nights, starting the same day as the reservation.
Help Desk : I'm sorry. It looks like they're full. Would you like me to look for something else?
Customer : Yes please. Is there something else available in that area?
Help Desk : I'm sorry, it looks like that is the only 3 star hotel in that area, would you like me to look somewhere else?
Customer : Can we try for 1 night instead of 2? I'll need the reference number please.
Help Desk : I was able to get one night, the reference number is 9XVT8M5T.
Customer : Thank you so much!

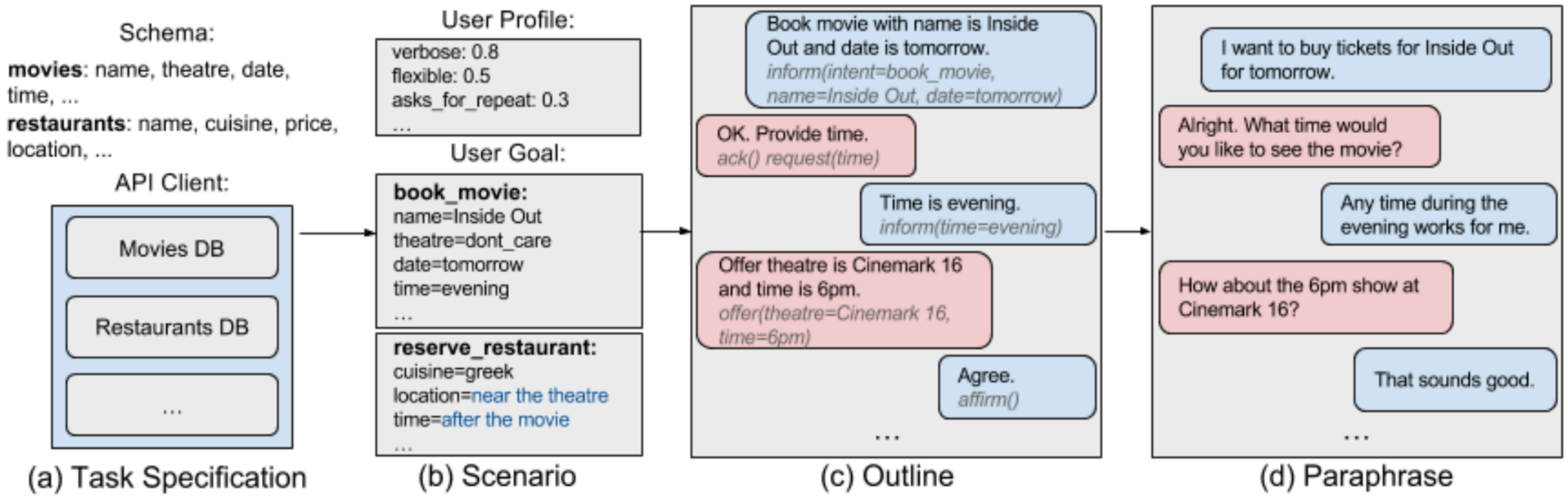
Please **modifiy** the following answers based on the latest customer response:

- What does the user want?

Is the user looking for a specific restaurant **by name**?	not mentioned
What is the **food type** the user wants?	indian
What is the **area** the user wants?	centre
What is the **price range** the user wants?	expensive

Wen et al, 2017

Example: M2M Data Collection



Schema:

movies: name, theatre, date, time, ...

restaurants: name, cuisine, price, location, ...

API Client:

(a) Task Specification

User Profile:

verbose: 0.8
flexible: 0.5
asks_for_repeat: 0.3
...

User Goal:

book_movie:
name=Inside Out
theatre=dont_care
date=tomorrow
time=evening
...

reserve_restaurant:
cuisine=greek
location=near the theatre
time=after the movie
...

(b) Scenario

Book movie with name is Inside Out and date is tomorrow.
inform(intent=book_movie, name=Inside Out, date=tomorrow)

OK. Provide time.
ack() request(time)

Time is evening.
inform(time=evening)

Offer theatre is Cinemark 16 and time is 6pm.
offer(theatre=Cinemark 16, time=6pm)

Agree.
affirm()

...

(c) Outline

I want to buy tickets for Inside Out for tomorrow.

Alright. What time would you like to see the movie?

Any time during the evening works for me.

How about the 6pm show at Cinemark 16?

That sounds good.

...

(d) Paraphrase

Shah et al, 2018

Short Summary

- Data is the lifeblood of machine learning algorithms.
- Historically, AI breakthroughs tend to correlate more to the datasets available in that area, rather than algorithms.
- In-domain conversational data is hard to get. Therefore, finding a way to collect/generate data is critical.
- PolyAI is building an in-house data collection platform to support client's business need.

Conclusion

Conclusion

- The dream of “*talking to machine*” is happening.
  - It is the matter how, not when.
- Existing tools do not provide the right kind of abstractions.
  - And there are so much to learn from the research community.
- A collective efforts are required to realize it
  - Across different industries.
  - Across multiple roles.
  - And multiple approaches.

Q & A

Interested in **partnership** or **working** at PolyAI?
Send email to: contact@poly-ai.com

POLYAi