Neural Approaches to Conversational AI

Jianfeng Gao, Michel Galley, Lihong Li

Microsoft Research, Google Inc.

SIGIR 2018

Ann Arbor Michigan, USA, July 8, 2018

Neural Approaches to Conversational AI

Jianfeng Gao, Michel Galley, Lihong Li

Microsoft Research, Google Inc.

ACL 2018

Melbourne, Australia, July 15, 2018

Contact Information:

Jianfeng Gao http://research.microsoft.com/~jfgao

Michel Galley http://research.microsoft.com/~mgalley

Lihong Li https://lihongli.github.io

Slides:

http://microsoft.com/en-us/research/publication/neural-approaches-to-conversational-ai/

Outline

- **Part 1: Introduction**
  - **Who should attend this tutorial**
  - **Dialogue: what kinds of problem**
  - **A unified view: dialogue as optimal decision making**
  - **Deep learning leads to paradigm shift in NLP and IR**
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- Part 4: Fully data-driven conversation models and chatbots

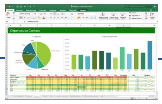
Who should attend this tutorial?

- NLP/IR community as our primary target audience
- Whoever wants to understand and create modern dialogue agents that
  - Can chat like a human
  - Can answer questions of various topics (movie stars, theory of relativity)
  - Can fulfill tasks (whether report, travel planning)
  - Can help make business decision
- Focus on neural approaches, but symbolic approaches are still widely used

Aspirational Goal: Enterprise Assistant

Where are sales lagging behind our forecast?

Task Completion



The worst region is [country], where sales are XX% below projections

Do you know why?

QA (decision support)

The forecast for [product] growth was overly optimistic

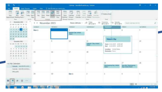
How can we turn this around?

Here are the 10 customers in [country] with the most growth potential, per our CRM model

Can you set up a meeting with the CTO of [company]?

Info Consumption

Task Completion



Yes, I've set up a meeting with [person name] for next month when you're in [location]

Thanks

What kinds of problems?

"I am smart"	Turing Test ("I" talk like a human)
"I have a question"	Information consumption
"I need to get this done"	Task completion
"What should I do?"	Decision support

What kinds of problems?

“I am smart”	Turing Test
“I have a question”	Information consumption
“I need to get this done”	Task completion
“What should I do?”	Decision support

- *What is the employee review schedule?*
- *What room is the project review meeting in?*
- *When is the ACL 2018 conference?*
- *What does DNN stand for?*

What kinds of problems?

"I am smart"	Turing Test
"I have a question"	Information consumption
"I need to get this done"	Task completion
"What should I do?"	Decision support

- *Book me the biz trip to San Francisco*
- *Reserve a table at Kisaku for 5 people, 7PM tonight*
- *Brief me on people in my Thursday 9:00 am meeting*
- *Schedule a meeting with Bill at 10:00 tomorrow.*

What kinds of problems?

"I am smart"	Turing Test
"I have a question"	Information consumption
"I need to get this done"	Task completion
"What should I do?"	Decision support

- *Why are sales in China so far behind forecast?*

What kinds of problems?

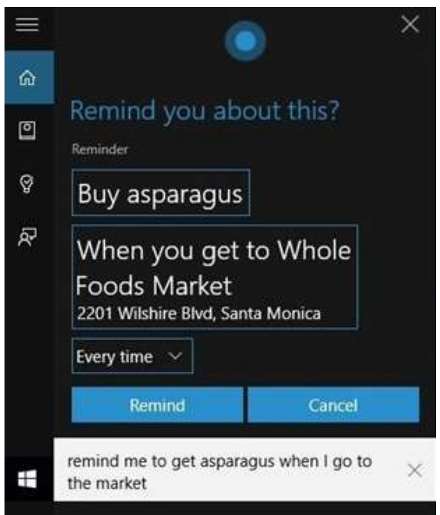
"I am smart"	Turing Test ("I" talk like a human)
"I have a question"	Information consumption
"I need to get this done"	Task completion
"What should I do?"	Decision support

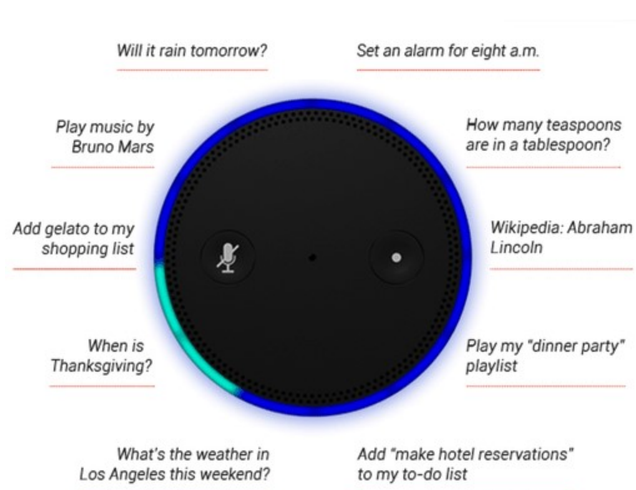
Chitchat (social bot): Turing Test ("I" talk like a human)

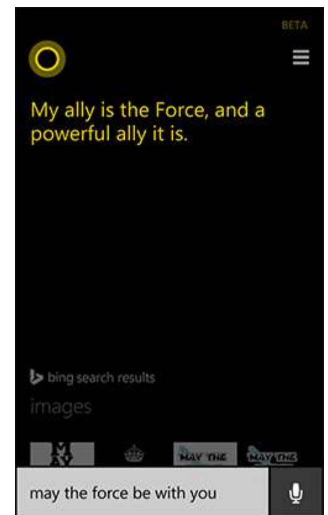
Goal-oriented dialogues: Information consumption, Task completion, Decision support

Personal assistants today

Google Now Siri Cortana amazon echo



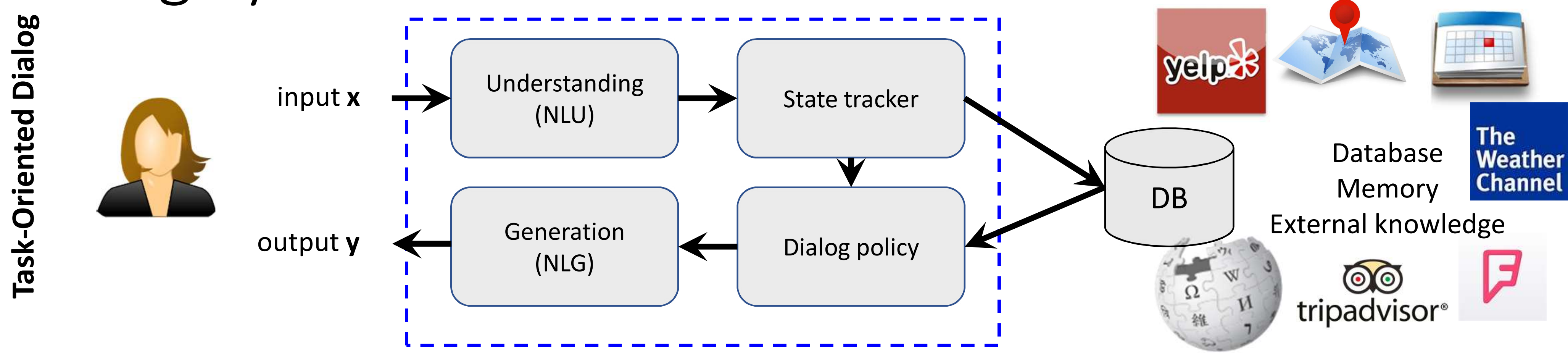


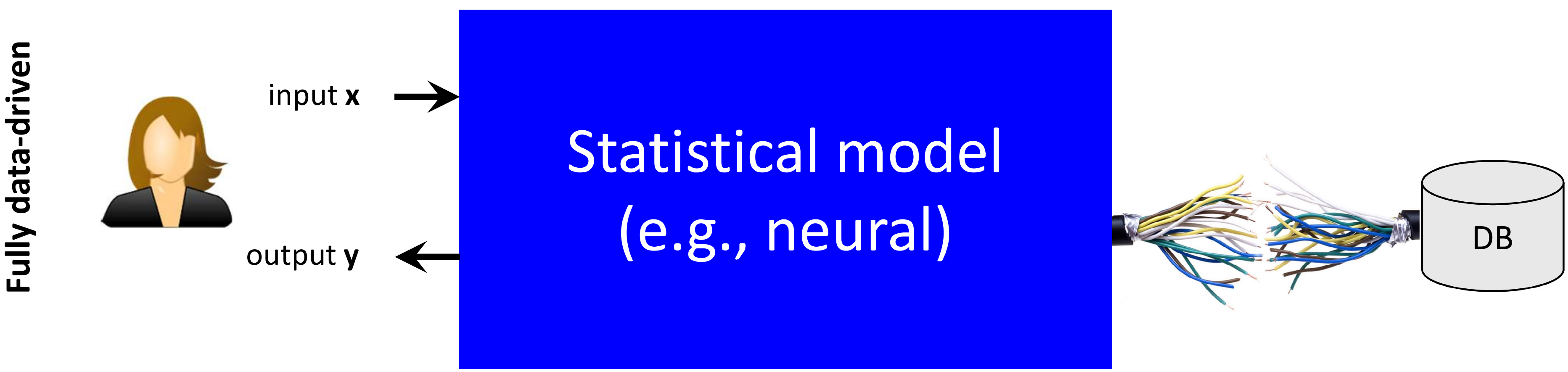


goal oriented

Engaging (social bots)

Dialog Systems





[Young+ 13; Tur & De Mori 11; Ritter+ 11; Sordoni+ 15; Vinyals & Le 15; Shang+ 15; etc.]

A unified view: dialogue as optimal decision making

- Dialogue as a Markov Decision Process (MDP)
  - Given state s, select action a according to (hierarchical) policy π
  - Receive reward r, observe new state a'
  - Continue the cycle until the episode terminates.
- Goal of dialogue learning: find optimal π to maximize expected rewards

Dialogue	State (s)	Action (a)	Reward (r)
Info Bots (Q&A bot over KB, Web etc.)	Understanding of user Intent (belief state)	Clarification questions, Answers	Relevance of answer # of turns
Task Completion Bots (Movies, Restaurants, …)	Understanding of user goal (belief state)	Dialog act + slot_value	Task success rate # of turns
Social Bot (XiaoIce)	Conversation history	Response	Engagement

Traditional definition of NLP: the branch of AI

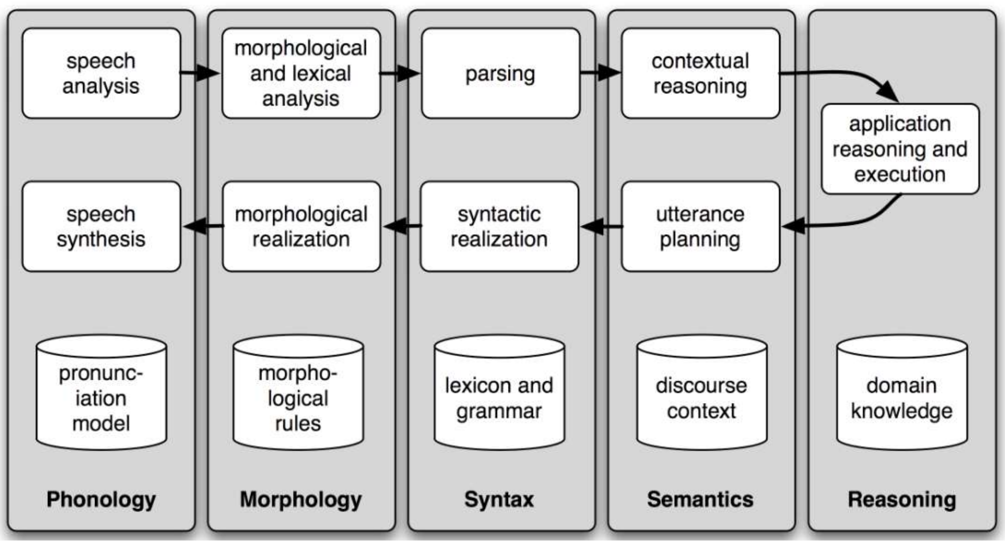
- Deal with analyzing, understanding and generating the languages that humans use naturally (natural language)
- Study *knowledge of language* at different levels
  - Phonetics and Phonology – the study of linguistic sounds
  - Morphology – the study of the meaning of components of words
  - Syntax – the study of the structural relationships between words
  - Semantics – the study of meaning
  - Discourse – they study of linguistic units larger than a single utterance

[Jurafsky & Martin 09]

Pragmatic definition: building computer systems

- Process large text corpora, turning information into knowledge
  - Text classification
  - Information retrieval and extraction
  - Machine reading comprehension and question answering
  - …
- Enable human-computer interactions, making knowledge accessible to humans in the most natural way
  - Dialogue and conversational agents
  - Machine translation
  - …

Traditional NLP component stack



1. **Natural language understand (NLU):** parsing (speech) input to semantic meaning and update the system state
2. **Application reasoning and execution:** take the next action based on state
3. **Natural language generation (NLG):** generating (speech) response from action

[Arul Menezes & Bill Dolan. 2017. The transition of natural language processing to neural methods. Unpublished report.]

Challenge of NLP: the diversity of natural language

Many-to-many mapping btw *symbolic* language and *semantic* meaning

Ambiguity

Example: I made her duck.

- *I cooked waterfowl for her.*
- *I cooked waterfowl belonging to her.*
- *I created the plaster duck she owns.*
- *I caused her to quickly lower her head or body.*
- *I waved my magic wand and turned her into undifferentiated waterfowl.*

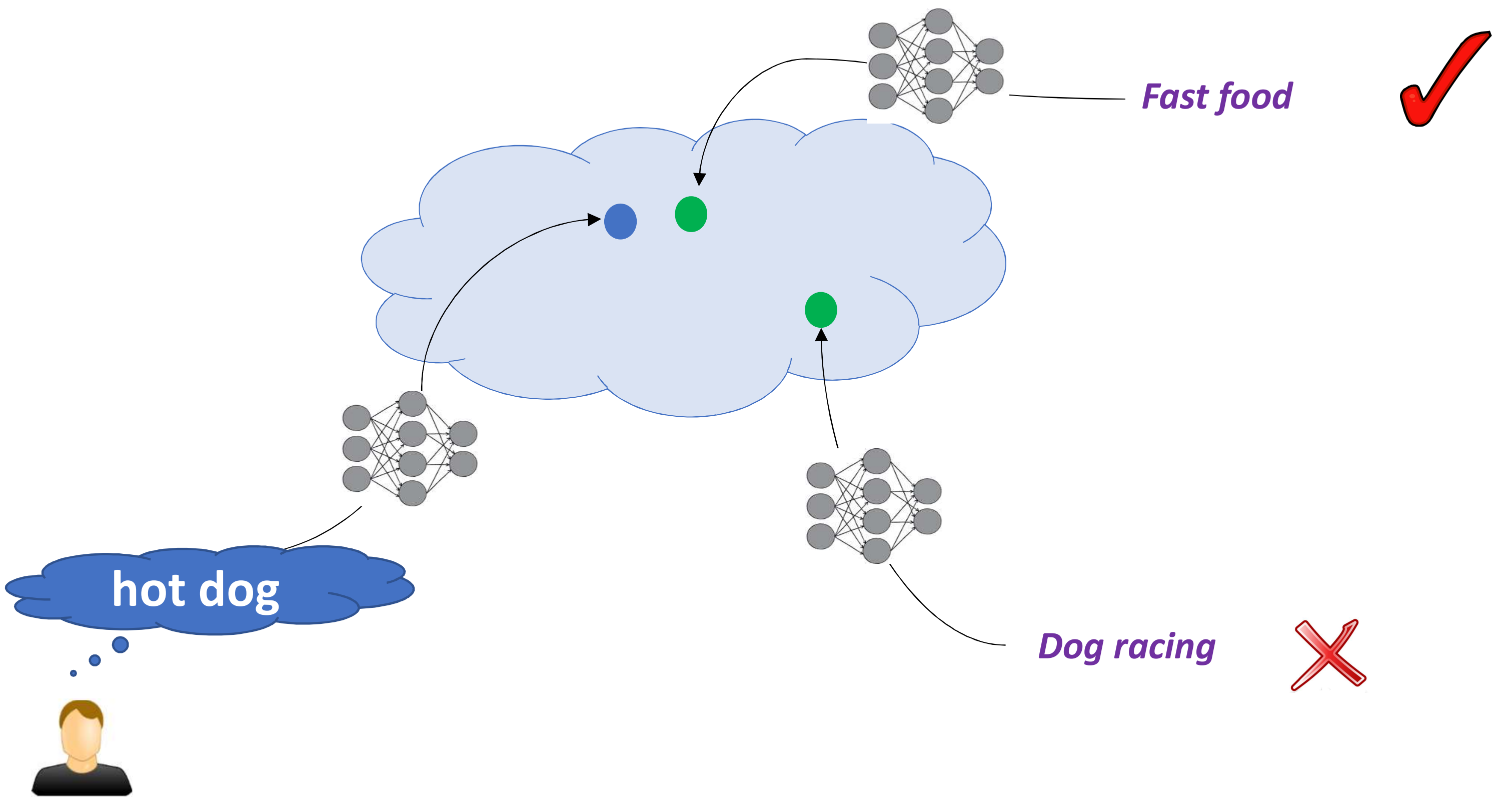
Paraphrase

Example: How long is the X river?

- *The Mississippi River is 3,734 km (2,320 mi) long.*
- *...is a short river, some 4.5 miles (7.2 km) in length*
- *The total length of the river is 2,145 kilometers.*
- *… at the estimated length of 5,464 km (3,395 mi)…*
- *… has a meander length of 444 miles (715 km)…*
- *… Bali's longest river, measuring approximately 75 kilometers from source to mouth.*
- *The … mainstem is 2.75 miles (4.43 km) long although total distance from headwater source tributaries to the sea is 14 miles (23 km).*

[Jurafsky & Martin 09; Dolan 17]

Mapping from symbolic to semantic *via DNN*?



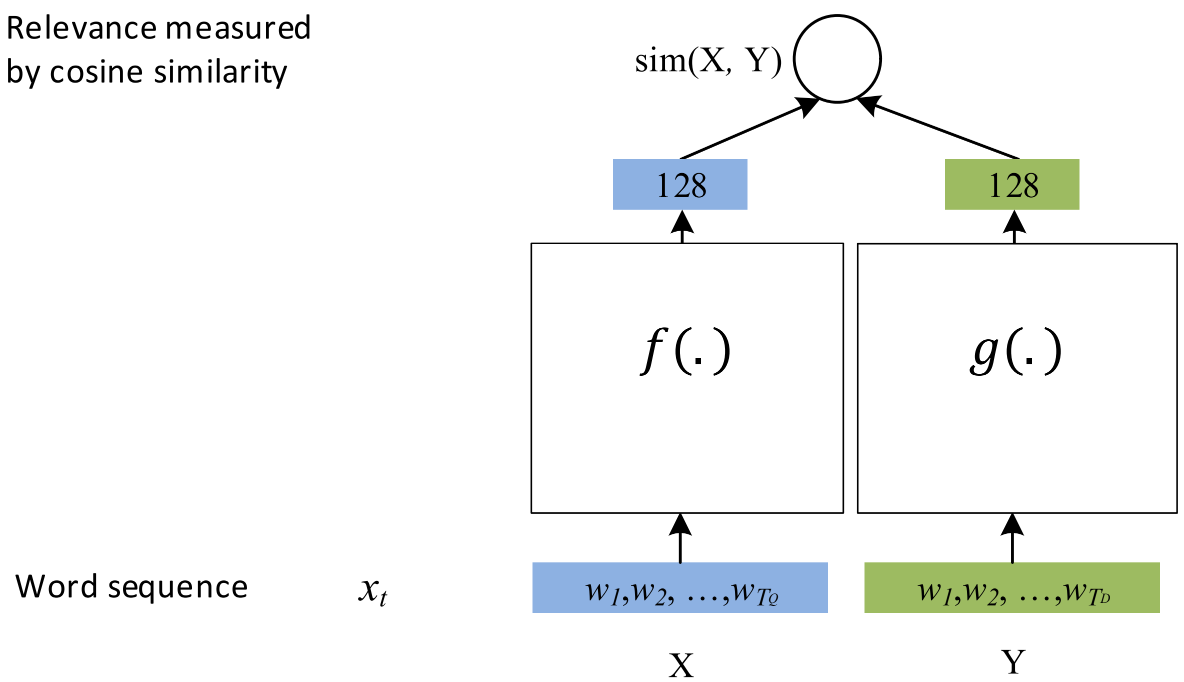
Deep Semantic Similarity Model (DSSM)

- Compute semantic similarity between two text strings X and Y
  - Map X and Y to feature vectors in a latent semantic space via deep neural net
  - Compute the cosine similarity between the feature vectors

Tasks	X	Y	Ref
Web search	*Search query*	*Web document*	Huang+ 13; Shen+ 14; Palangi+ 16
Entity linking	*Entity mention and context*	*Entity and its corresponding page*	Gao+ 14b
Online recommendation	*Doc in reading*	*Interesting things / other docs*	Gao+ 14b
Image captioning	*Image*	*Text*	Fang+ 15
Machine translation	*Sentence in language A*	*Translations in language B*	Gao+ 14a
Question answering	*Question*	*Answer*	Yih+ 15

Sent2Vec (DSSM) http://aka.ms/sent2vec

DSSM: Compute Similarity in Semantic Space



Learning: maximize the similarity between X (source) and Y (target)

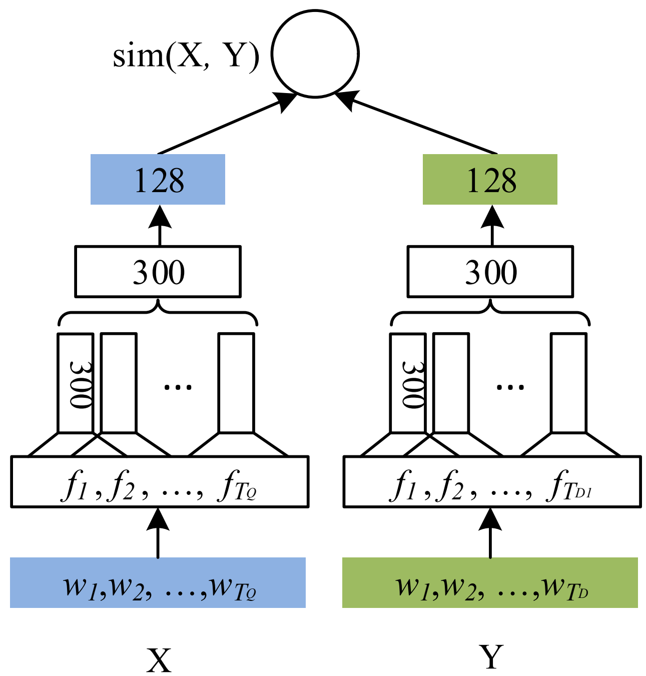
Representation: use DNN to extract abstract semantic features, f or g is a

- Multi-Layer Perceptron (MLP) if text is a bag of words [Huang+ 13]
- **Convolutional Neural Network (CNN) if text is a bag of chunks** [Shen+ 14]
- Recurrent Neural Network (RNN) if text is a sequence of words [Palangi+ 16]

DSSM: Compute Similarity in Semantic Space

Relevance measured by cosine similarity

Semantic layer	h
Max pooling layer	v
Convolutional layer	c_t
Word hashing layer	f_t
Word sequence	x_t



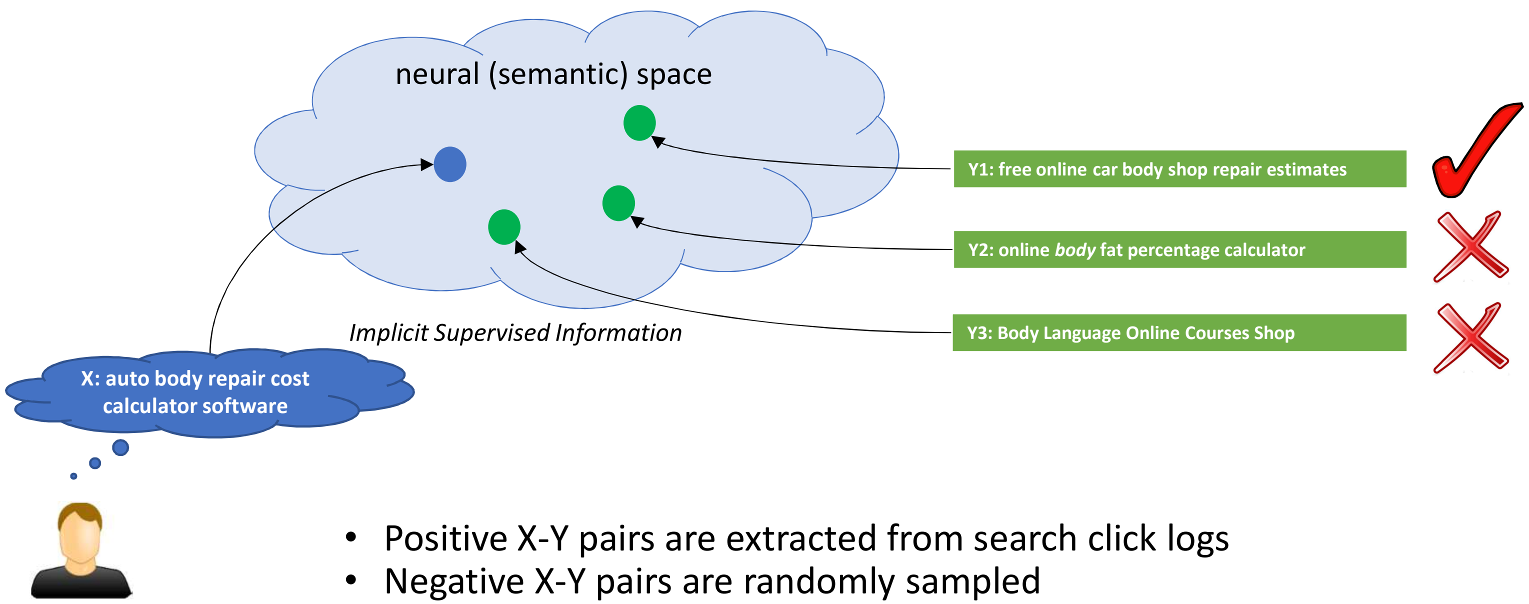
Learning: maximize the similarity between X (source) and Y (target)

Representation: use DNN to extract abstract semantic representations

Convolutional and Max-pooling layer: identify key words/concepts in X and Y

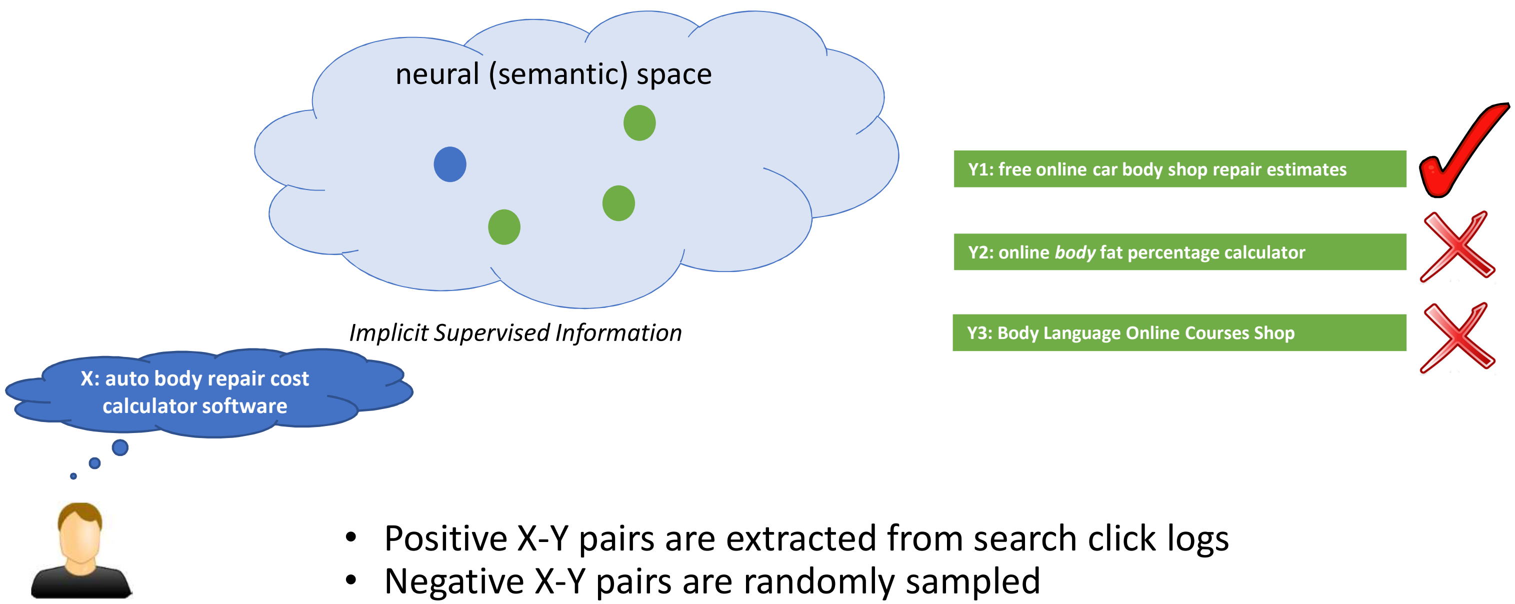
Word hashing: use sub-word unit (e.g., letter n-gram) as raw input to handle very large vocabulary

Learning DSSM from Labeled X-Y Pairs



- Positive X-Y pairs are extracted from search click logs
- Negative X-Y pairs are randomly sampled
- Map X and Y into the same semantic space via deep neural net

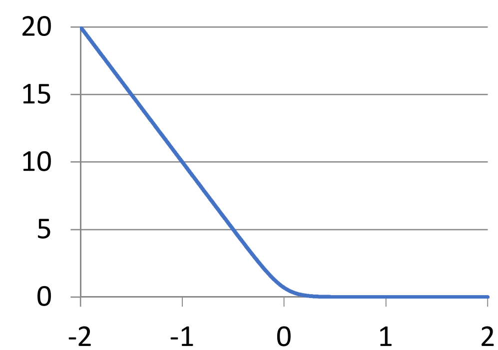
Learning DSSM from Labeled X-Y Pairs



- Positive X-Y pairs are extracted from search click logs
- Negative X-Y pairs are randomly sampled
- Map X and Y into the same semantic space via deep neural net
- Positive Y are closer to X than negative Y in that space

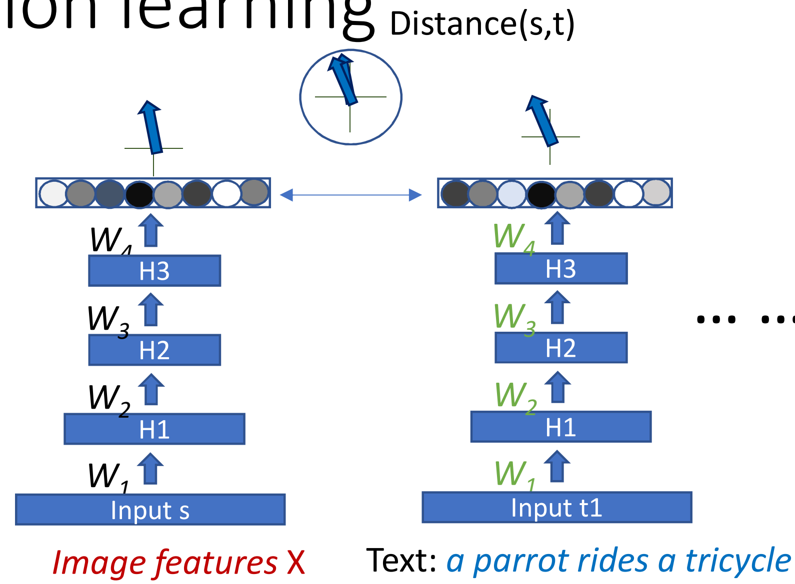
Learning DSSM from Labeled X-Y Pairs

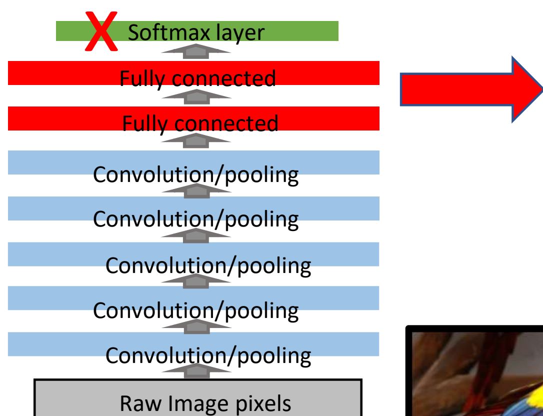
- Consider a query X and two docs Y^+ and Y^-
  - Assume Y^+ is more relevant than Y^- with respect to X
- $\text{sim}_{\boldsymbol{\theta}}(X, Y)$ is the cosine similarity of X and Y in semantic space, mapped by DSSM parameterized by $\boldsymbol{\theta}$
- $\Delta = \text{sim}_{\boldsymbol{\theta}}(X, Y^+) - \text{sim}_{\boldsymbol{\theta}}(X, Y^-)$
  - We want to maximize Δ
- $Loss(\Delta; \boldsymbol{\theta}) = \log(1 + \exp(-\gamma\Delta))$
- Optimize $\boldsymbol{\theta}$ using mini-batch SGD on GPU



Go beyond text:
DSSM for multi-modal representation learning

- Recall DSSM for text input pairs: (X, Y)
- Now: replace text X by image **X**
- Using DNN/CNN features of image
- Can rank/generate text given image or can rank images given text.





[Fang+ 15]

Paradigm shift in NLP/IR

1. From symbolic to neural computation via semantic representation learning
    - Due to novel DNN architectures and learning algorithms; leads to high accuracy in many tasks
2. New Applications and Experience
    - E.g., link language to real-world signals such as images and machine state
3. End-to-end Learning
    - Simplifies systems, reduces effort for feature engineering and localization
4. Deep Reinforcement Learning
    - Makes it possible to build intelligent agents for real-world applications such as goal-oriented dialogue

[Menezes & Dolan 17]

Outline

- Part 1: Introduction
- **Part 2: Question answering (QA) and machine reading comprehension (MRC)**
  - **Knowledge base QA tasks**
  - **From semantic parsing to embedding-based approaches**
  - **Multi-turn knowledge base QA agents**
  - **Neural MRC models for text-based QA**
- Part 3: Task-oriented dialogue
- Part 4: Fully data-driven conversation models and chatbots

Open-Domain Question Answering (QA)

Will I qualify for OSAP if I'm new in Canada?

Selected Passages from Bing

"Visit the OSAP website for application deadlines. To get OSAP, you have to be eligible. You can apply using an online form, or you can print off the application forms. If you submit a paper application, you must pay an application fee. The online application is free."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/how-do-i-apply-for-the-ontario-student-assistance-program-osap/

"To be eligible to apply for financial assistance from the Ontario Student Assistance Program (OSAP), you must be a: 1 Canadian citizen; 2 Permanent resident; or 3 Protected person/convention refugee with a Protected Persons Status Document (PPSD)."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/who-is-eligible-for-the-ontario-student-assistance-program-osap/

"You will not be eligible for a Canada-Ontario Integrated Student Loan, but can apply for a part-time loan through the Canada Student Loans program. There are also grants, bursaries and scholarships available for both full-time and part-time students."

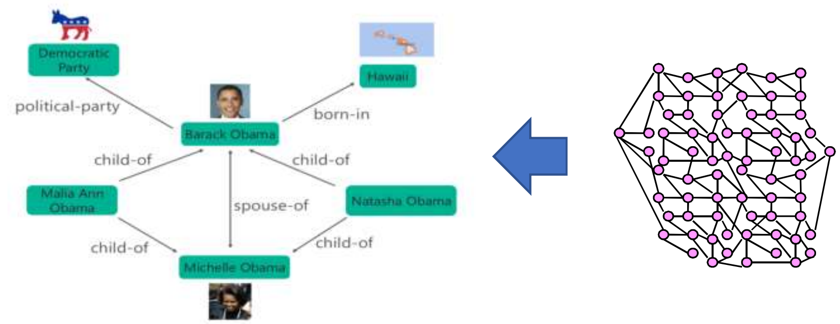
Source: http://www.campusaccess.com/financial-aid/osap.html

Answer
No. You won't qualify.

Text-QA

What is Obama's citizenship?

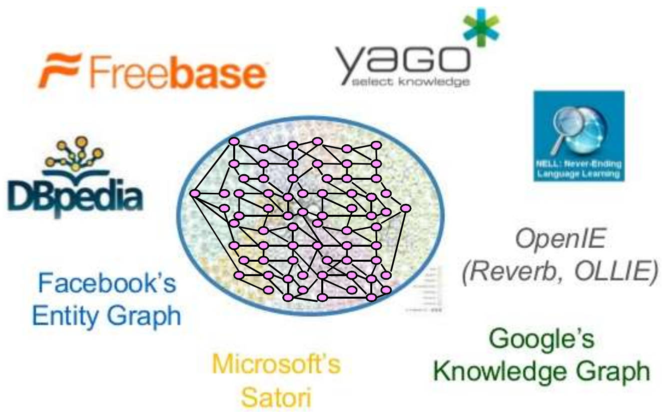
Selected subgraph from Microsoft's Satori

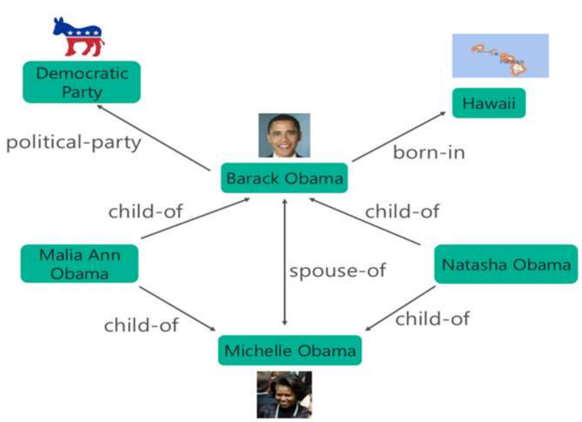


Answer
USA

Knowledge Base (KB)-QA

Question Answering (QA) on Knowledge Base





Large-scale knowledge graphs

- Properties of billions of entities
- Plus relations among them

An QA Example:

Question: `what is Obama's citizenship?`

- Query parsing:
  `(Obama, Citizenship,?)`
- Identify and infer over relevant subgraphs:
  `(Obama, BornIn, Hawaii)`
  `(Hawaii, PartOf, USA)`
- correlating semantically relevant relations:
  `BornIn ~ Citizenship`

Answer: `USA`

Symbolic approaches to QA

- Understand the question via **semantic parsing**
  - Input: what is Obama's citizenship?
  - Output (LF): (Obama, Citizenship,?)
- Collect relevant information via fuzzy **keyword matching**
  - (Obama, BornIn, Hawaii)
  - (Hawaii, PartOf, USA)
  - Needs to know that BornIn and Citizenship are semantically related
- Generate the answer via **reasoning**
  - (Obama, Citizenship, **USA**)
- **Challenges**
  - **Paraphrasing in NL**
  - **Search complexity of a big KG**

[Richardson+ 98; Berant+ 13; Yao+ 15; Bao+ 14; Yih+ 15; etc.]

Example: “How long is the X river?”

- The Mississippi River is 3,734 km (2,320 mi) long.
- …is nearly 86 km long…
- ...is a short river, some 4.5 miles (7.2 km) in length
- The total length of the river is 2,145 kilometres (1,333 mi).
- … at the estimated length of 5,464 km (3,395 mi)…
- **…**is a 25-mile (40 km) tributary of …
- … has a meander length of 444 miles (715 km)…
- … Bali’s longest river, measuring approximately 75 kilometers from source to mouth.
- The … mainstem is 2.75 miles (4.43 km) long although total distance from headwater source tributaries to the sea is 14 miles (23 km).

- …is 314 km long
- …is nearly 86 km long…
- … is a 92-mile (148 km) long tributary of the…
- ...is a short river, some 4.5 miles (7.2 km) in length
- …flows nearly 20 miles (32 km) to the west
- The [river], which is 6,853 km (4,258 miles) long…
- It runs a course of about 105 kilometers
- The 1,450-mile-long (2,330 km) [river] drains…
- ...a 234-mile (377-kilometer) man-made waterway…
- … at the estimated length of 5,464 km (3,395 mi)…
- ... stretches for 2,639 miles (4,247 km).
- …is a 25-mile (40 km) tributary of …
- …starting in and flowing for nearly 160 kilometers through….
- …flows almost 70 stream miles.
- The river runs 184 kilometers before joining…
- … Bali's longest river, measuring approximately 75 kilometers from source to mouth.
- …is reported to be anywhere from 5,499 to 6,690 kilometres (3,417 to 4,157 mi). Often it is said to be "about" 6,650 kilometres (4,130 mi) long.
- ...reaches a length of approximately 25 kilometres
- The length of the Ouse alone is about 52 miles (84 km).
- Measuring a length of 60 kilometers, the [river] flows through
- It has a total length of 925 km (575 mi).
- The total length of the river is 2,145 kilometres (1,333 mi).
- Its length is 209 km…
- …is about 1,180 miles (1,900 km) in length.
- ...the river flows for more than 1,200 km (750 mi)
- …the river proper flows only for 113 km…
- …flows slowly for 900 kilometres (560 mi)…
- … has a meander length of 444 miles (715 km)…
- …is a 350-kilometre (220 mi) long river in …
- it …meanders slowly southwards for 2,320 miles (3,730 km) to …
- The river's main stem is about 71 miles (114 km) long. Its length to its most distant headwater tributary is about 220 miles (350 km).
- After approximately 30 kilometres (19 mi) of its 78-kilometre (48 mi) course, it ….
- …is the longest river in the United Kingdom, at about 220 miles (354 km).
- … is the second-longest river in Central and Western Europe (after the Danube), at about 1,230 km (760 mi)…
- The … mainstem is 2.75 miles (4.43 km) long although total distance from headwater source tributaries to the sea is 14 miles (23 km).
- At 320 kilometres (200 mi) (with some estimates ranging up to 596 kilometres (370 mi))...

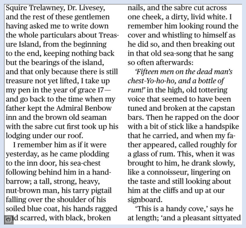
Key Challenge in KB-QA – Language Mismatch

- Lots of ways to ask the same question
  - *“What was the date that Minnesota became a state?”*
  - *“Minnesota became a state on?”*
  - *“When was the state Minnesota created?”*
  - *“Minnesota's date it entered the union?”*
  - *“When was Minnesota established as a state?”*
  - *“What day did Minnesota officially become a state?”*

- Need to map them to the predicate defined in KB
  - location.dated_location.date_founded

Symbolic Space

- **Knowledge Representation**
  - *explicitly* stored as words, relations, templates
  - high-dim, discrete, sparse vectors
- **Inference**
  - *slow* on a big knowledge base
  - keyword matching is sensitive to paraphrase alternations
- **Human comprehensible but not computationally efficient**

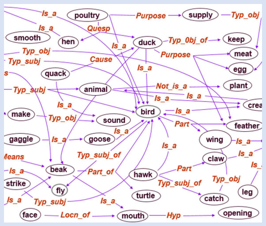


Squire Trelawney, Dr. Livesey, and the rest of these gentlemen having asked me to write down the whole particulars about Treasure Island, from the beginning to the end, keeping nothing back but the bearings of the island, and that only because there is still treasure not yet lifted, I take up my pen in the year of grace 17— and go back to the time when my father kept the Admiral Benbow inn and the brown old seaman with the sabre cut first took up his lodging under our roof.

I remember him as if it were yesterday, as he came plodding to the inn door, his sea-chest following behind him in a hand-barrow; a tall, strong, heavy, nut-brown man, his tarry pigtail falling over the shoulder of his soiled blue coat, his hands ragged and scarred, with black, broken nails, and the sabre cut across one cheek, a dirty, livid white. I remember him looking round the cover and whistling to himself as he did so, and then breaking out in that old sea-song that he sang so often afterwards:

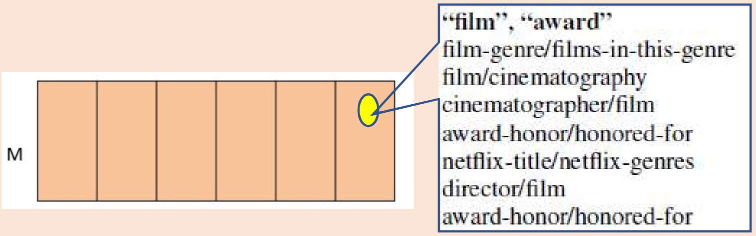
'Fifteen men on the dead man's chest-Yo-ho-ho, and a bottle of rum!' in the high, old tottering voice that seemed to have been tuned and broken at the capstan bars. Then he rapped on the door with a bit of stick like a handspike that he carried, and when my father appeared, called roughly for a glass of rum. This, when it was brought to him, he drank slowly, like a connoisseur, lingering on the taste and still looking about him at the cliffs and up at our signboard.

'This is a handy cove,' says he at length; 'and a pleasant sittyated



Neural Space

- **Knowledge Representation**
  - *Implicitly* stored as semantic concepts/classes
  - low-dim, cont., dense vectors
- **Inference**
  - *fast* on compact memory
  - semantic matching is robust to paraphrase alternations
- **Computationally efficient but not human comprehensible *yet***

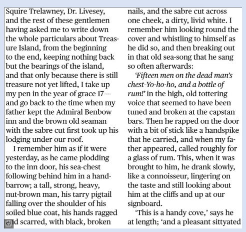


From symbolic to neural computation

Input: Q

Symbolic → Neural
by **Encoding** (Q/D/Knowledge)

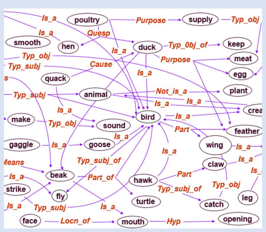
Symbolic Space
- human readable


Squire Trelawney, Dr. Livesey, and the rest of these gentlemen having asked me to write down the whole particulars about Treasure Island, from the beginning to the end, keeping nothing back but the bearings of the island, and that only because there is still treasure not yet lifted, I take up my pen in the year of grace 17— and go back to the time when my father kept the Admiral Benbow inn and the brown old seaman with the sabre cut first took up his lodging under our roof.

I remember him as if it were yesterday, as he came plodding to the inn door, his sea-chest following behind him in a hand-barrow; a tall, strong, heavy, nut-brown man, his tarry pigtail falling over the shoulder of his soiled blue coat, his hands ragged and scarred, with black, broken nails, and the sabre cut across one cheek, a dirty, livid white. I remember him looking round the cover and whistling to himself as he did so, and then breaking out in that old sea-song that he sang so often afterwards:

'Fifteen men on the dead man's chest-Yo-ho-ho, and a bottle of rum!' in the high, old tottering voice that seemed to have been tuned and broken at the capstan bars. Then he rapped on the door with a bit of stick like a handspike that he carried, and when my father appeared, called roughly for a glass of rum. This, when it was brought to him, he drank slowly, like a connoisseur, lingering on the taste and still looking about him at the cliffs and up at our signboard.

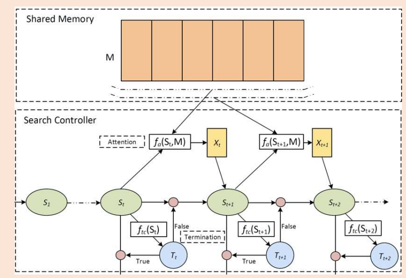
'This is a handy cove,' says he at length; 'and a pleasant sittyated



Error(A, A*)

Reasoning: Question + KB → answer vector via multi-step inference, summarization, deduction etc.

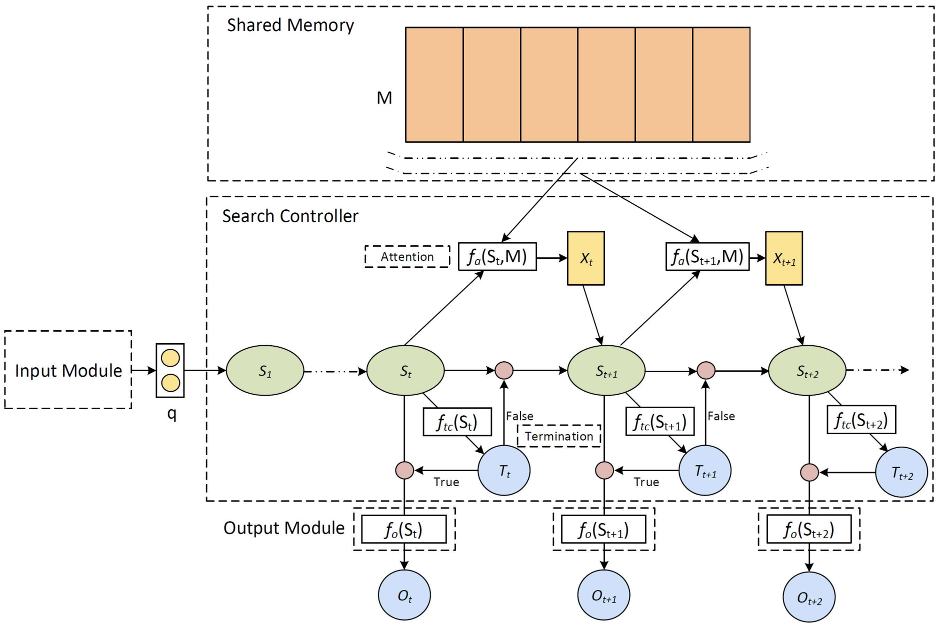
Neural Space
- Computationally efficient



Output: A

Neural → Symbolic
by **Decoding** (synthesizing answer)

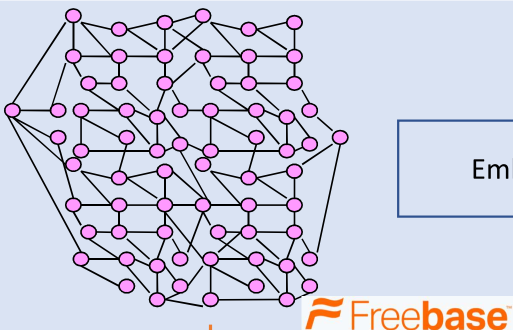
Case study: ReasoNet with Shared Memory



- **Shared memory** (M) encodes task-specific knowledge
  - **Long-term memory:** encode KB for answering all questions in QA on KB
  - **Short-term memory:** encode the passage(s) which contains the answer of a question in QA on Text
- **Working memory** (hidden state S_t) contains a description of the current state of the world in a reasoning process
- **Search controller** performs multi-step inference to update S_t of a question using knowledge in shared memory
- Input/output modules are task-specific

[Shen+ 16; Shen+ 17]

Joint learning of Shared Memory and Search Controller

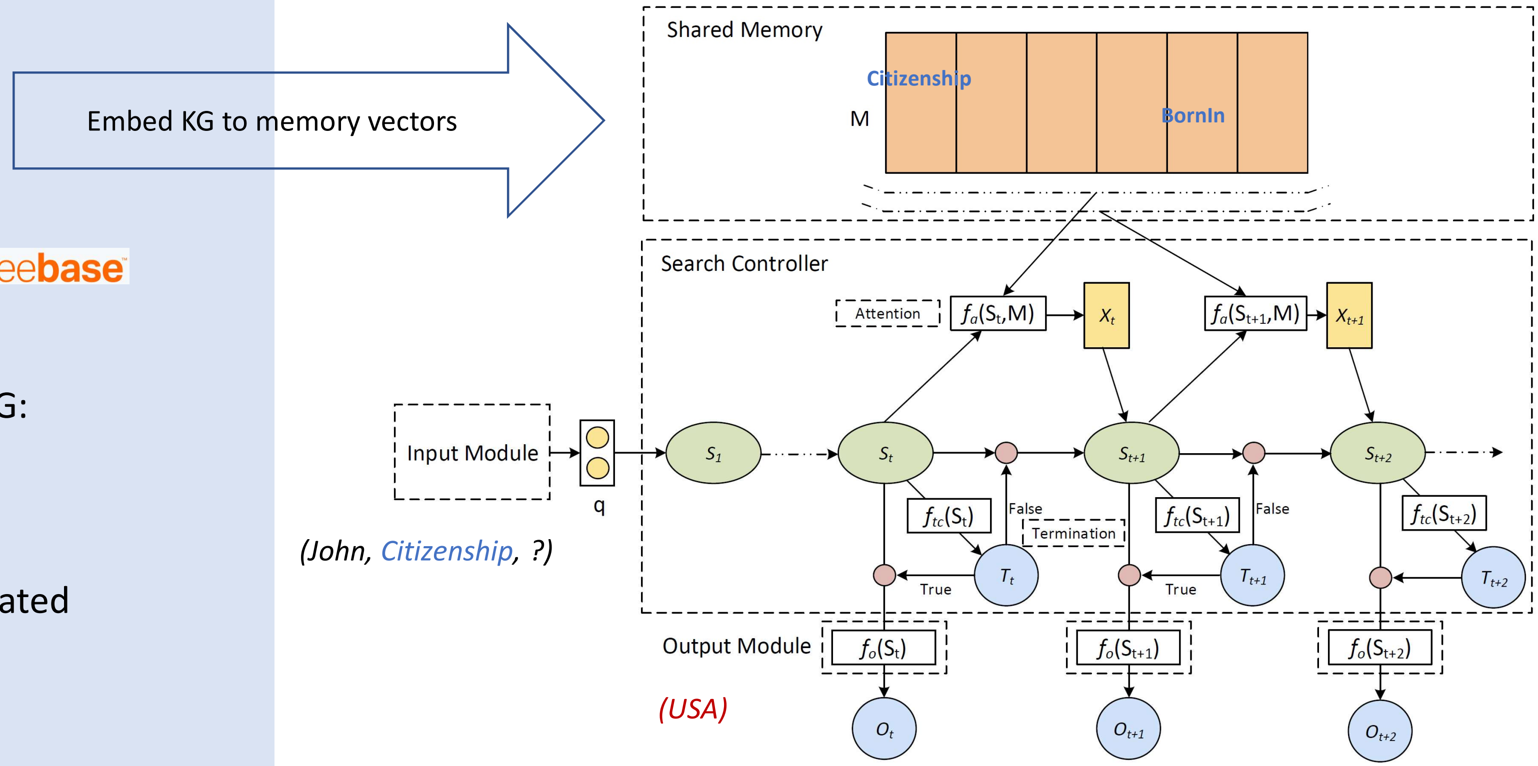


Paths extracted from KG:

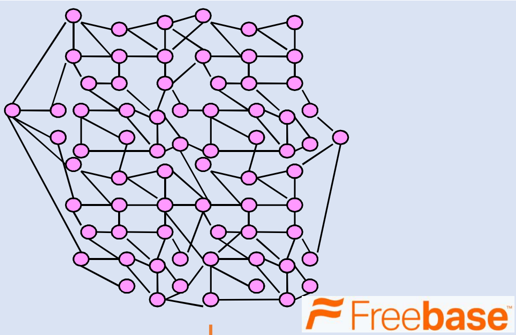
(John, BornIn, Hawaii)
(Hawaii, PartOf, USA)
(John, *Citizenship*, USA)
…

Training samples generated

(John, BornIn, ?)->(Hawaii)
(Hawaii, PartOf, ?)->(USA)
(John, Citizenship, ?)->(USA)
…



Joint learning of Shared Memory and Search Controller

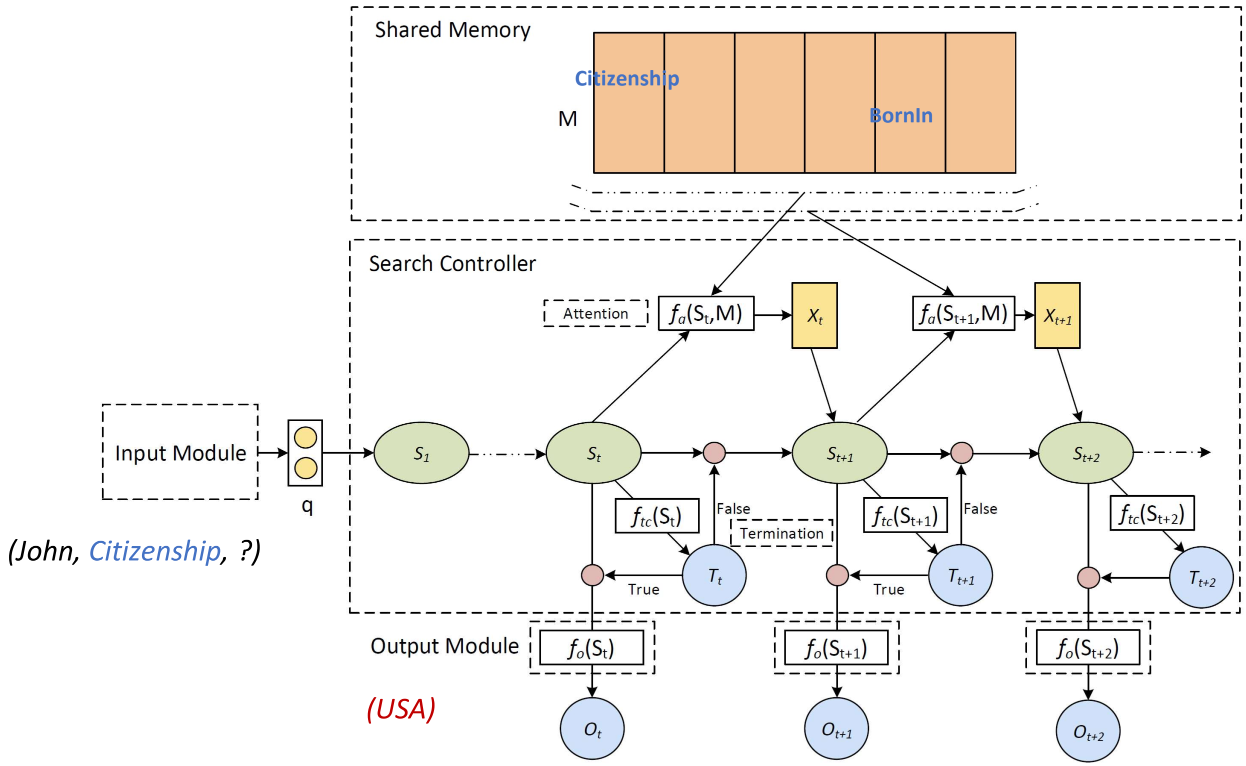


Paths extracted from KG:

(John, BornIn, Hawaii)
(Hawaii, PartOf, USA)
(John, *Citizenship*, USA)
…

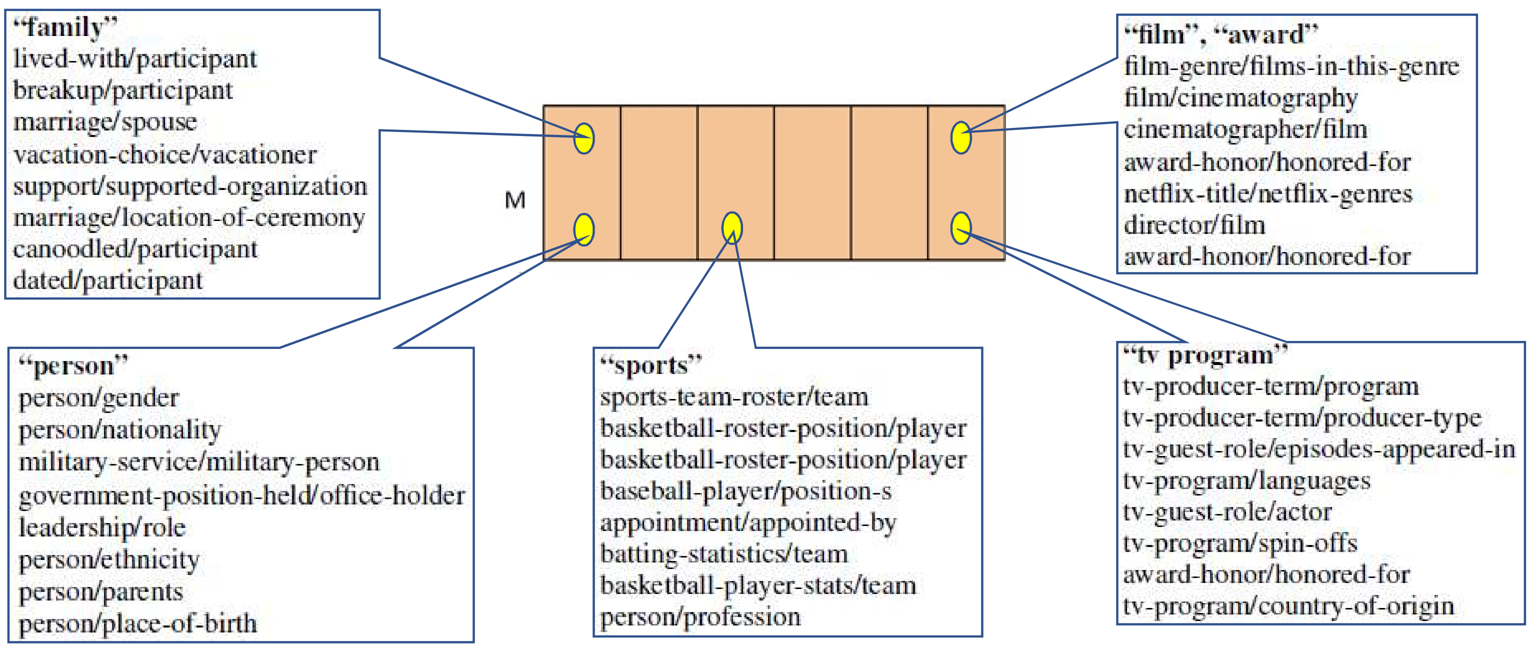
Training samples generated

(John, BornIn, ?)->(Hawaii)
(Hawaii, PartOf, ?)->(USA)
(John, Citizenship, ?)->(USA)
…

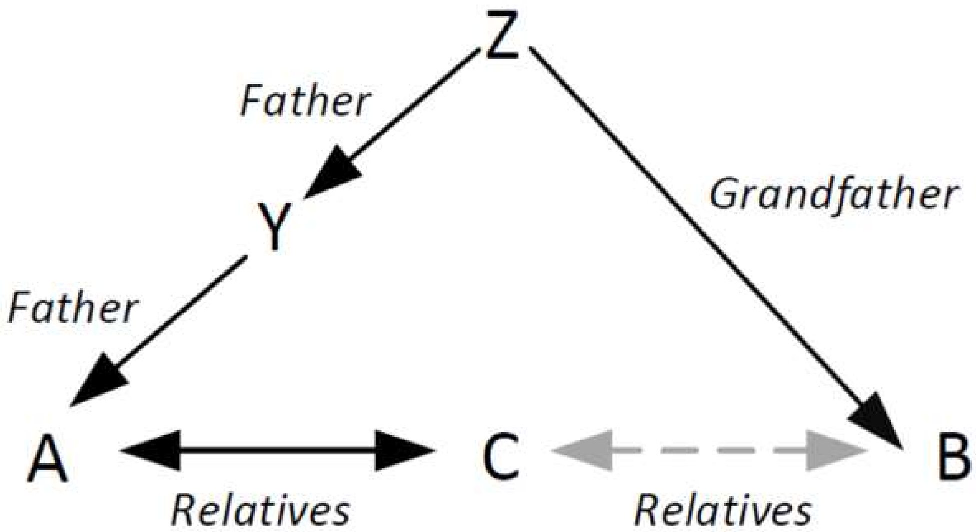


Shared Memory: long-term memory to store learned knowledge, like human brain

- Knowledge is learned via performing tasks, e.g., update memory to answer new questions
- New knowledge is *implicitly* stored in memory cells via gradient update
- Semantically relevant relations/entities can be compactly represented using similar vectors.



Search controller for KB QA



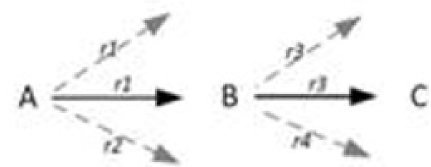
Space	Inference Path
Symbolic	$B \xrightarrow{\text{Grandfather}^{-1}} Z \xrightarrow{\text{Father}} Y \xrightarrow{\text{Father}} A \xrightarrow{\text{Relatives}} C$
Neural	$B \xrightarrow{\text{Relatives}} ? \Rightarrow A \xrightarrow{\text{Relatives}} C$

[Shen+ 16]

Reasoning over KG in symbolic vs neural spaces

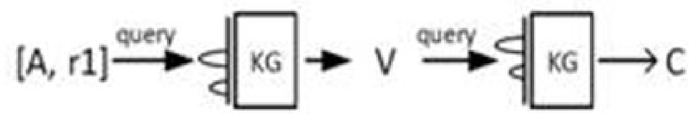
Symbolic: comprehensible but not robust

- Development: writing/learning production rules
- Runtime : random walk in symbolic space
- E.g., PRA [Lao+ 11], MindNet [Richardson+ 98]



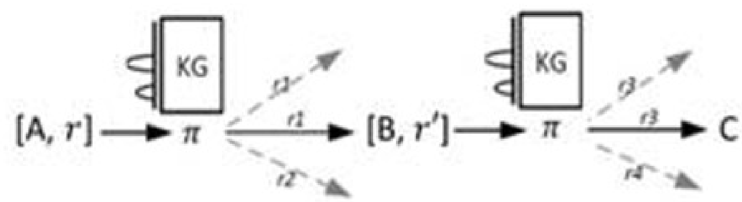
Neural: robust but not comprehensible

- Development: encoding knowledge in neural space
- Runtime : multi-turn querying in neural space (similar to nearest neighbor)
- E.g., ReasoNet [Shen+ 16a], DistMult [Yang+ 15]



Hybrid: robust and comprehensible

- Development: learning policy π that maps states in neural space to actions in symbolic space via RL
- Runtime : graph walk in symbolic space guided by π
- E.g., M-Walk [Shen+ 18], DeepPath [Xiong+ 18], MINERVA [Das+ 18]



Multi-turn KB-QA: what to ask?

- Allow users to query KB interactively without composing complicated queries
- Dialogue policy (what to ask) can be
  - Programmed [Wu+ 15]
  - Trained via RL [Wen+ 16; Dhingra+ 17]

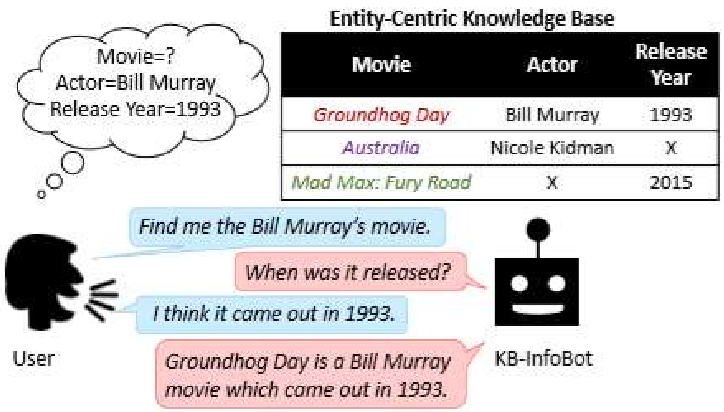


Figure 1: An interaction between a user looking for a movie and the KB-InfoBot. An entity-centric knowledge base is shown above the KB-InfoBot

Text-QA

Will I qualify for OSAP if I'm new in Canada?

Selected Passages from Bing

"Visit the OSAP website for application deadlines. To get OSAP, you have to be eligible. You can apply using an online form, or you can print off the application forms. If you submit a paper application, you must pay an application fee. The online application is free."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/how-do-i-apply-for-the-ontario-student-assistance-program-osap/

"To be eligible to apply for financial assistance from the Ontario Student Assistance Program (OSAP), you must be a: 1 Canadian citizen; 2 Permanent resident; or 3 Protected person/convention refugee with a Protected Persons Status Document (PPSD)."

Source: http://settlement.org/ontario/education/colleges-universities-and-institutes/financial-assistance-for-post-secondary-education/who-is-eligible-for-the-ontario-student-assistance-program-osap/

"You will not be eligible for a Canada-Ontario Integrated Student Loan, but can apply for a part-time loan through the Canada Student Loans program. There are also grants, bursaries and scholarships available for both full-time and part-time students."

Source: http://www.campusaccess.com/financial-aid/osap.html

Answer

No. You won't qualify.

MS MARCO [Nguyen+ 16]

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

What causes precipitation to fall?
gravity

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

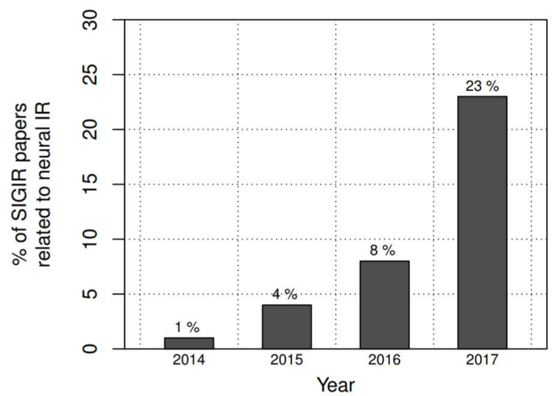
Where do water droplets collide with ice crystals to form precipitation?
within a cloud

Figure 1: Question-answer pairs for a sample passage in the SQuAD dataset. Each of the answers is a segment of text from the passage.

SQuAD [Rajpurkar+ 16]

MSMARCO Passage Ranking Challenge

MS MARCO



(src: http://bit.ly/fntir-neural)

Growing popularity of deep learning in IR, but lack of large training datasets

What is the "ImageNet" of ad-hoc retrieval?

New passage ranking task

- Collection size: 5-10M passages
- # of queries: 400-500K

Manual relevance annotations for "passage contains the answer"

Two task modes:

- Retrieve relevant passages from the full collection
- Re-rank top 1K candidates from BM25

Releasing soon (tentatively, **September 2018**)

Neural MRC Models

Input: Q

Symbolic → Neural
by **Encoding** (Q/D/Knowledge)

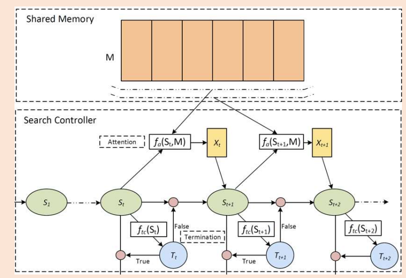
Symbolic Space
- human readable

Where did Tesla live for much of his life?

Tesla was renowned for his achievements and showmanship, eventually earning him a reputation in popular culture as an archetypal "mad scientist". His patents earned him a considerable amount of money, much of which was used to finance his own projects with varying degrees of success.:121,154 He lived most of his life in a series of New York hotels, through his retirement. Tesla died on 7 January 1943. His work fell into relative obscurity after his death, but in 1960 the General Conference on Weights and Measures named the SI unit of magnetic flux density the tesla in his honor. There has been a resurgence in popular interest in Tesla since the 1990s.

Reasoning: Q + D → answer vector
via multi-step inference,
summarization, deduction etc.

Neural Space
- Computationally efficient



Output: A

Neural → Symbolic
by **Decoding** (synthesizing answer)

Examples: BiDAF and SAN

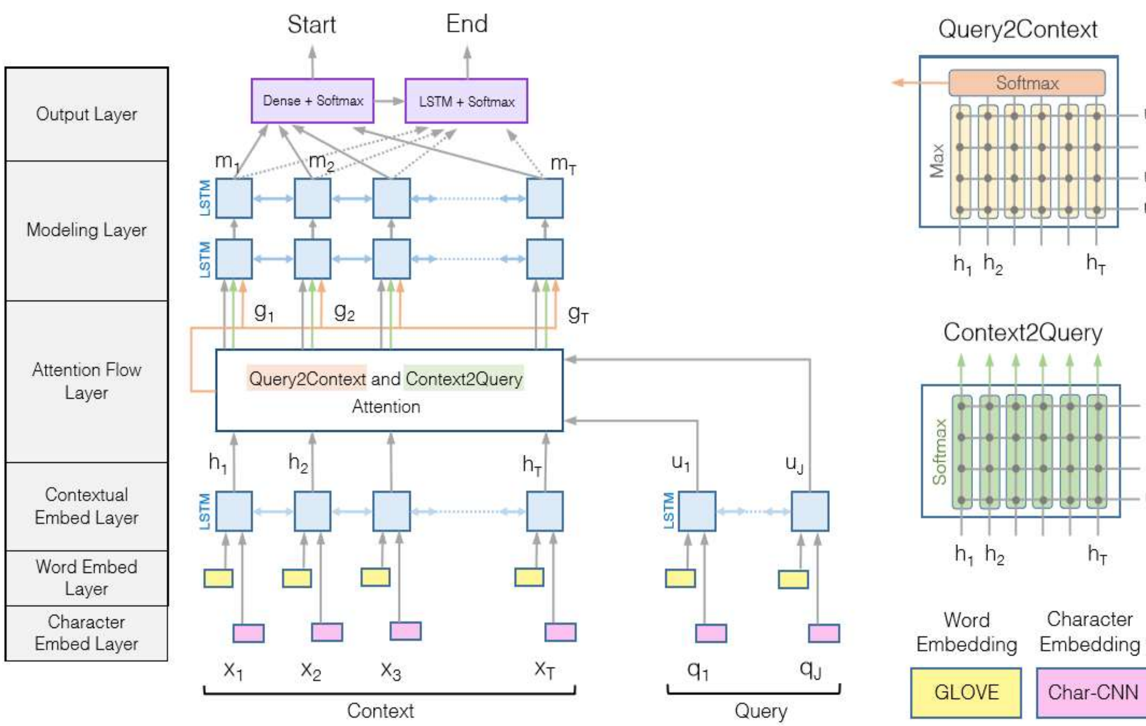
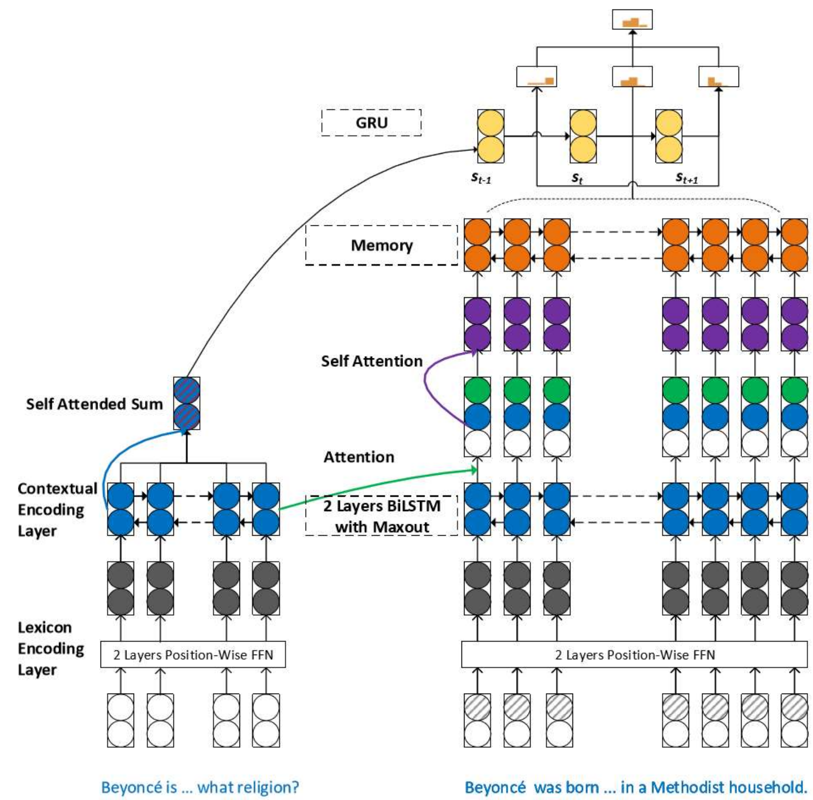


Figure 1: BiDirectional Attention Flow Model *(best viewed in color)*



[Seo+ 16; Liu+ 18]

Neural MRC Models on SQuAD

What types of European groups were able to avoid the plague?

> From Italy, the disease spread northwest across Europe, striking France, Spain, Portugal and England by June 1348, then turned and spread east through Germany and Scandinavia from 1348 to 1350. It was introduced in Norway in 1349 when a ship landed at Askøy, then spread to Bjørgvin (modern Bergen) and Iceland. Finally it spread to northwestern Russia in 1351. The plague was somewhat less common in parts of Europe that had smaller trade relations with their neighbours, including the Kingdom of Poland, the majority of the Basque Country, isolated parts of Belgium and the Netherlands, and isolated alpine villages throughout the continent.

A limited form of comprehension:

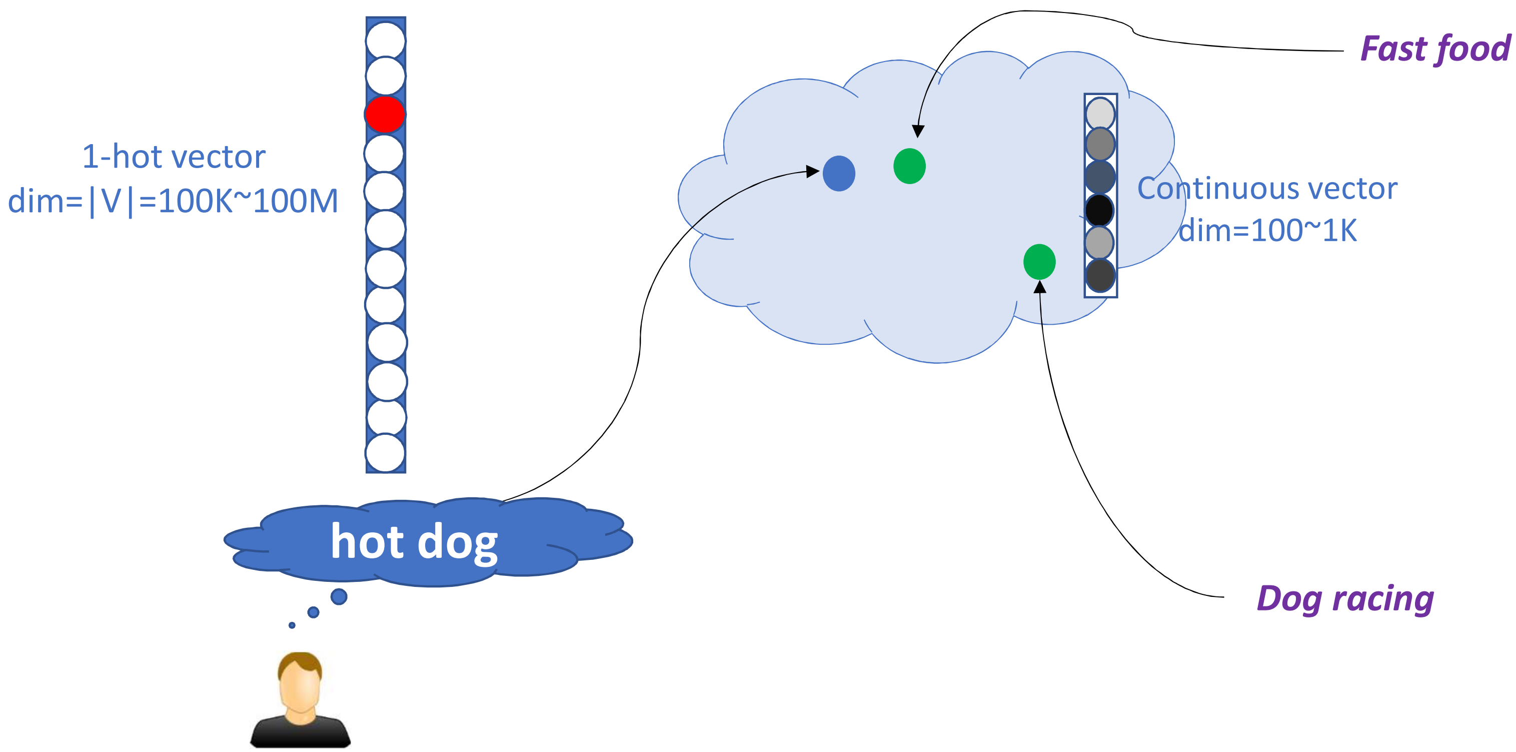
- No need for extra knowledge outside the paragraph
- No need for clarifying questions
- The answer must exist in the paragraph
- The answer must be a text span, not synthesized

- Encoding: map each text span to a semantic vector
- Reasoning: rank and re-rank semantic vectors
- Decoding: map the top-ranked vector to text

Three encoding components

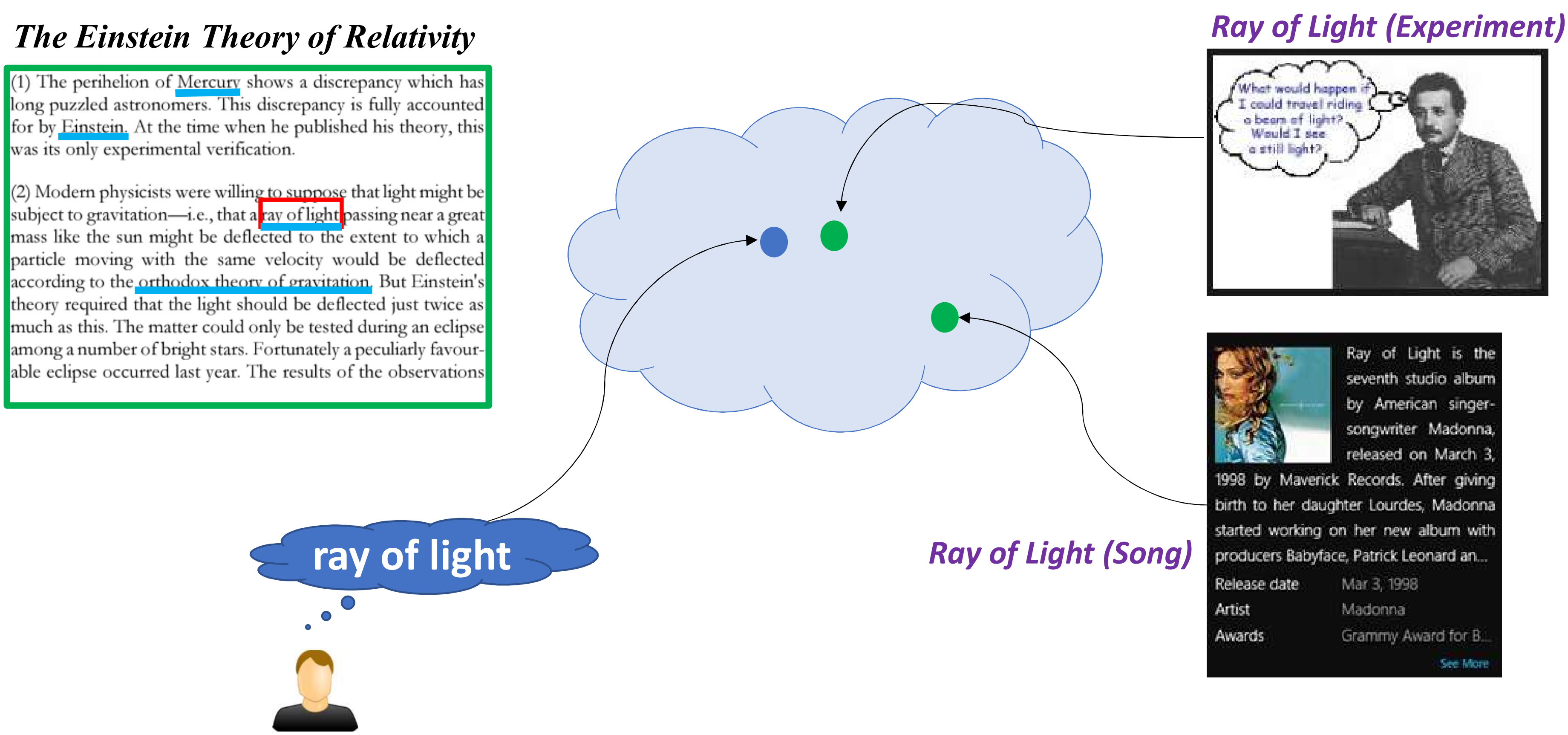
- Word embedding – word semantic space
  - represent each word as a low-dim continuous vector via GloVe [Pennington+ 14]
- Context embedding – contextual semantic space
  - capture context info for each word, via
    - BiLSTM [Melamud+ 16]
    - ELMo [Peter+ 18]: a task-specific combo of the intermediate layer representations of biLM
    - Stacked embedding encoder blocks [Yu+ 18]: convolution + self-attention
- Context-query attention – query dependent semantic space
  - fuse query info into passage via Attention
  - [Huang+ 17; Wang+ 17; Hu+ 17; Seo+ 16; Wang&Jiang 16]

Word embedding: word semantic space



[Mikolov+ 13; Pennington+ 14]

Context embedding: contextual semantic space

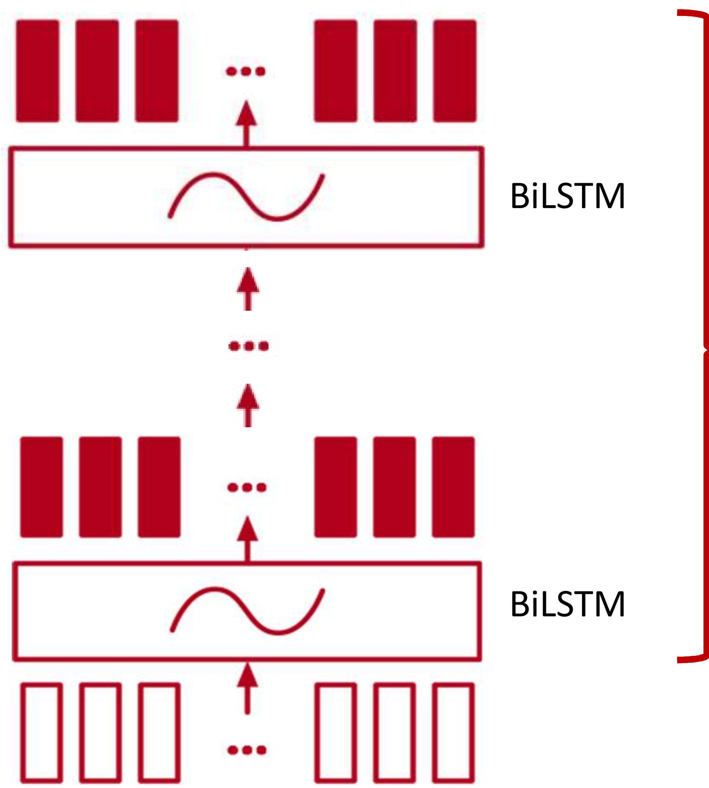


Context embedding via BiLSTM / ELMo

Context vectors $h_{t,L}$ at high level
One for each word with its context

Context vectors $h_{t,1}$ at low level
One for each word with its context

Embedding vectors x_t
One for each word

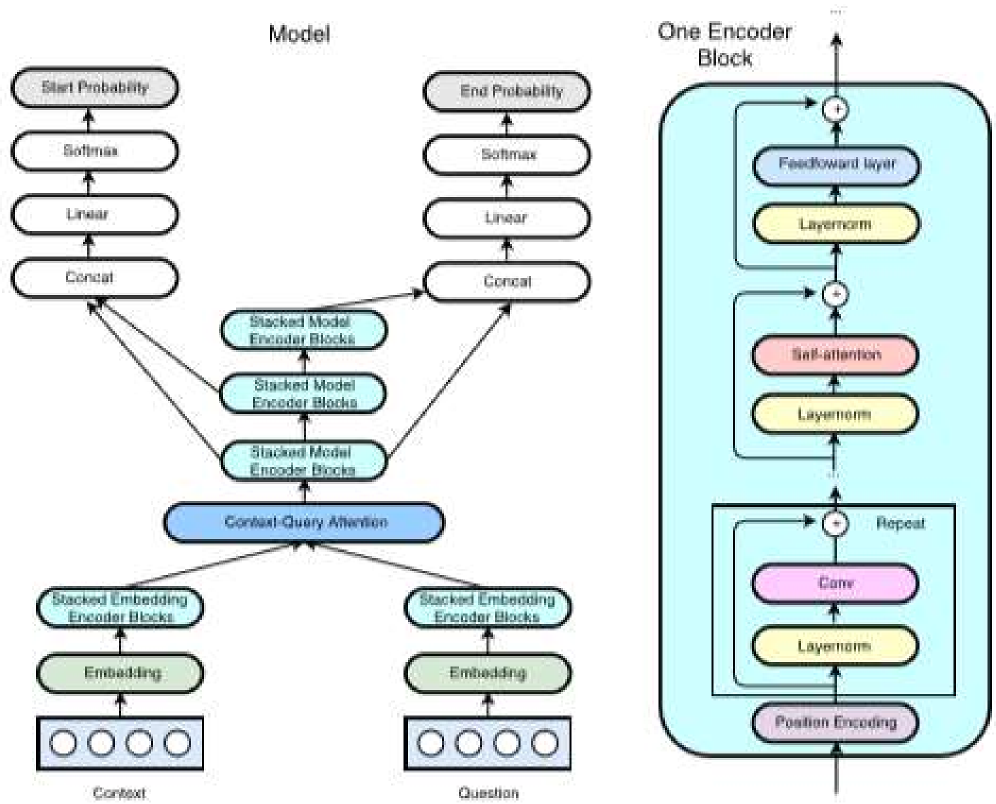


$$\text{ELMo}_t^{task} = \gamma^{task} \sum_{l=1 \ldots L} w_l^{task} h_{t,l}$$

Task-specific combination of hidden layers in BiLSTM

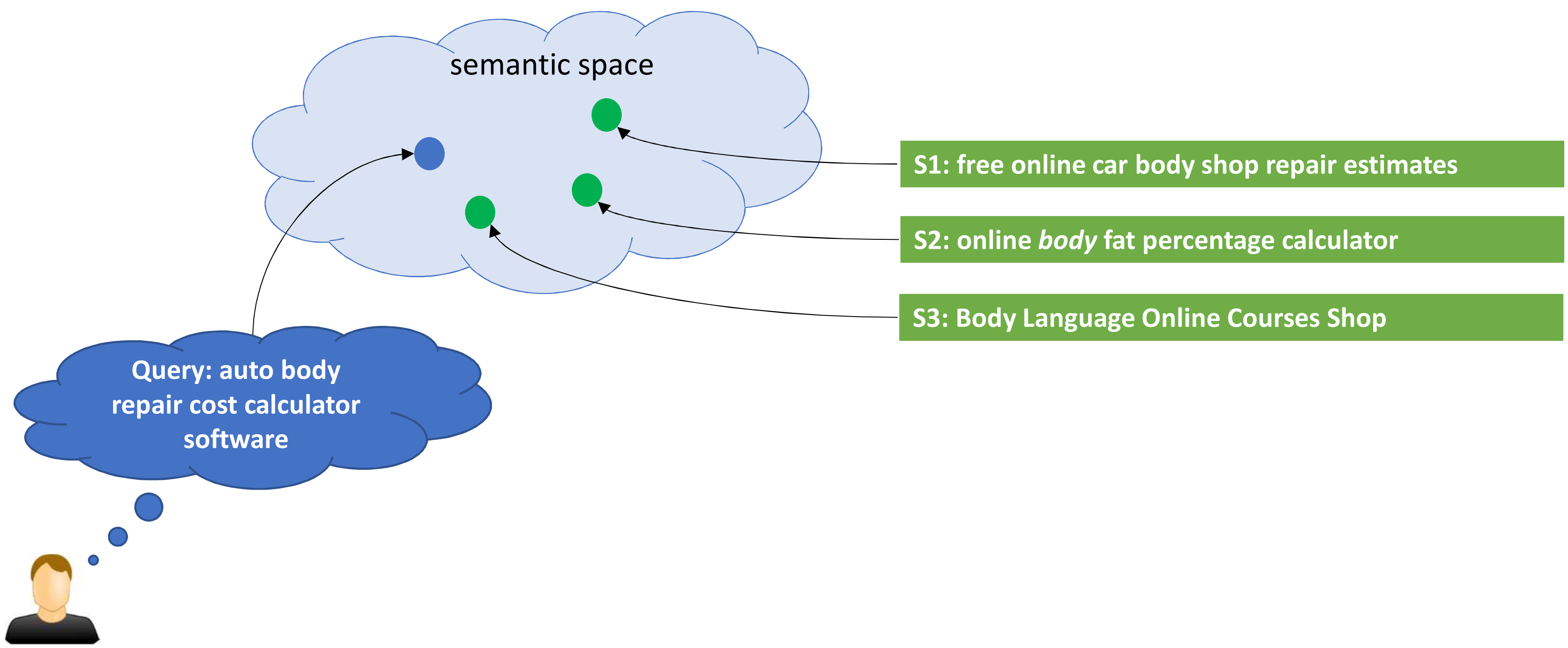
[Peter+ 18; McCann+ 17; Melamud+ 16]

Context embedding via self-attention and convolution

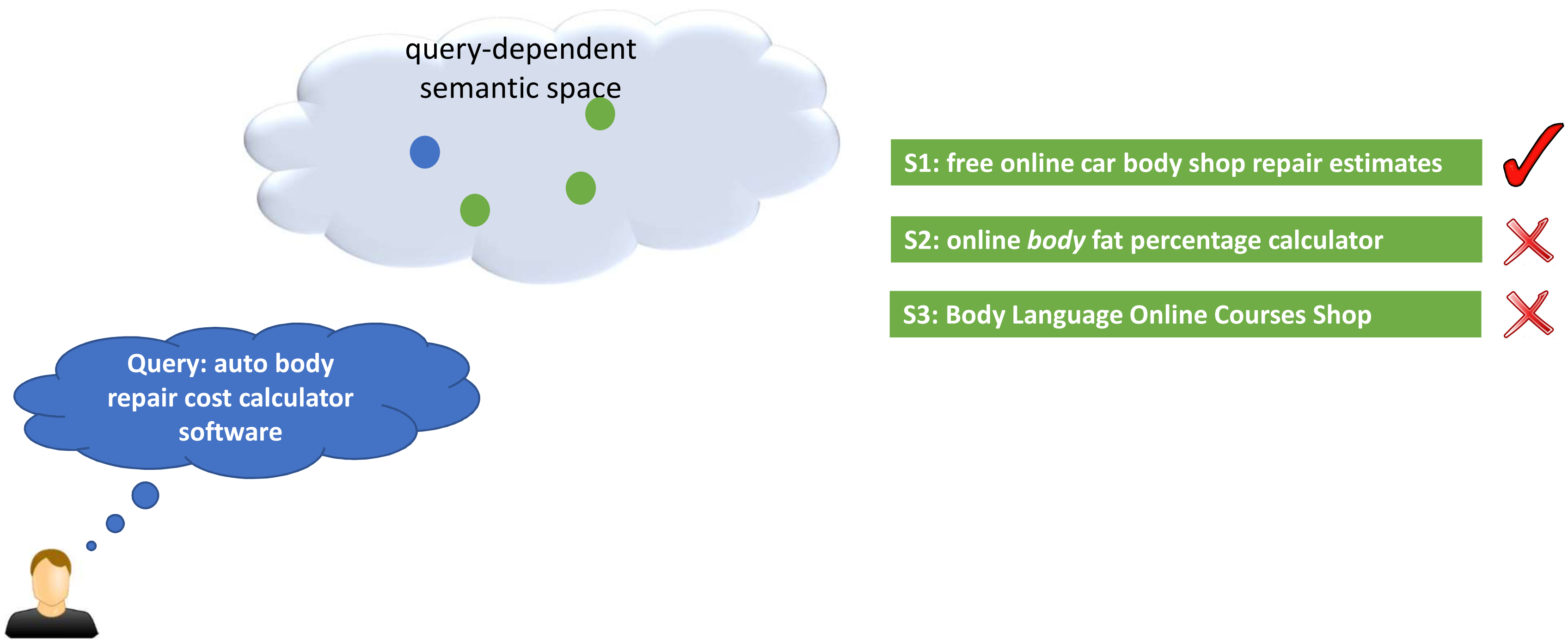


[Yu+ 18]

Context-query attention: query-dependent semantic space

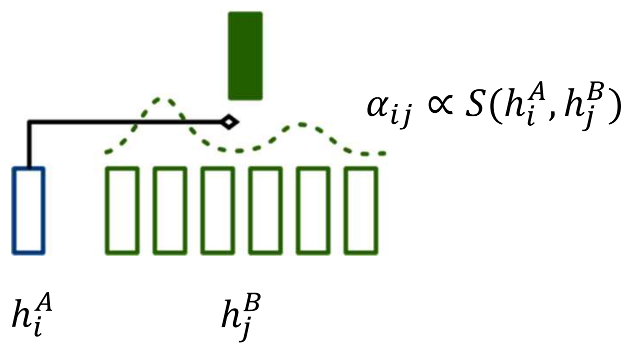


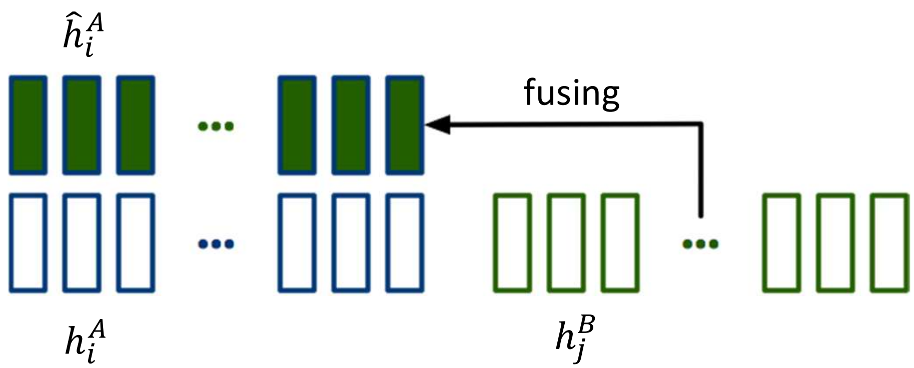
Context-query attention: query-dependent semantic space



Context-query attention: query-dependent semantic space

- Compute attention scores (similarity btw i and j): $S_{ij} = S(h_i^A, h_j^B)$
- Compute attention weights thru softmax: $\alpha_{ij} = \exp(S_{ij}) / \sum_k \exp(S_{ik})$
- Fusion info from B to A: $\hat{h}_i^A = \sum_j \alpha_{ij} h_j^B$





[Huang+ 17; Wang+ 17; Hu+ 17; Seo+ 16; Wang&Jiang 16]

Multi-step reasoning for Text-QA

- Learning to stop reading: dynamic multi-step inference
- Step size is determined based on the complexity of instance (QA pair)

Query	Who was the 2015 NFL MVP?
Passage	The Panthers finished the regular season with a 15–1 record, and quarterback **Cam Newton** was named the 2015 NFL Most Valuable Player (MVP).
Answer (1-step)	Cam Newton

Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and **Von Miller** for Denver.
Answer (3-step)	Von Miller

Multi-step reasoning: example

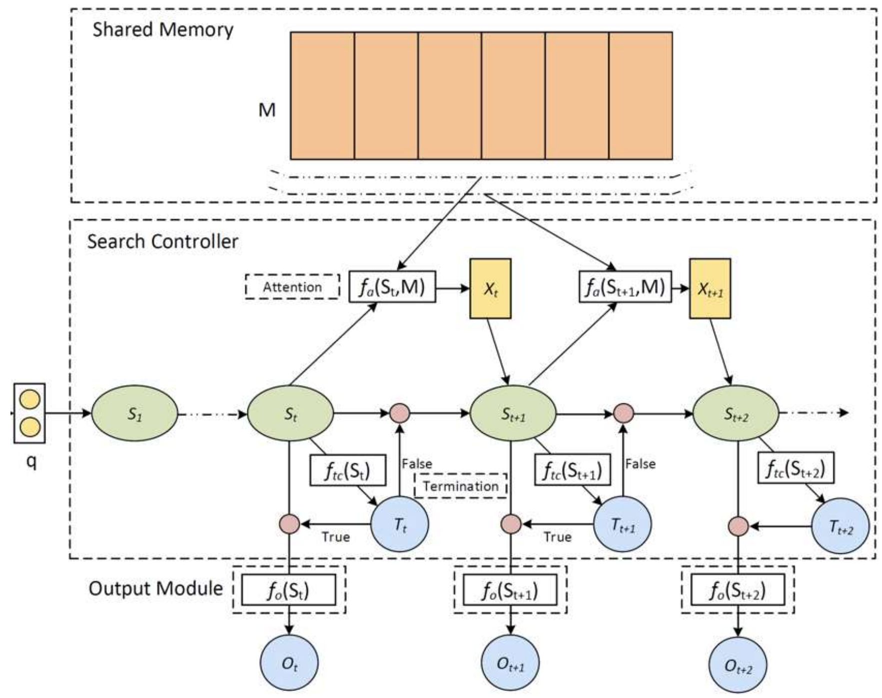
Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and **Von Miller** for Denver.
Answer	**Von Miller**

- Step 1:
  - **Extract:** Manning is #1 pick of 1998
  - **Infer:** Manning is NOT the answer
- Step 2:
  - **Extract:** Newton is #1 pick of 2011
  - **Infer:** Newton is NOT the answer
- Step 3:
  - **Extract:** Newton and Von Miller are top 2 picks of 2011
  - **Infer:** **Von Miller** is the #2 pick of 2011

ReasoNet: learn to stop reading

With Q in mind, read Doc repeatedly, each time focusing on different parts of doc until a satisfied answer is formed:

1. Given a set of docs in memory: $\mathbf{M}$
2. Start with query: S
3. Identify info in $\mathbf{M}$ that is related to S : $X = f_a(S, \mathbf{M})$
4. Update internal state: $S = \text{RNN}(S, X)$
5. Whether a satisfied answer O can be formed based on S: $f_{tc}(S)$
6. If so, stop and output answer $O = f_o(S)$; otherwise return to 3.



The step size is determined dynamically based on the complexity of the problem using reinforcement learning.

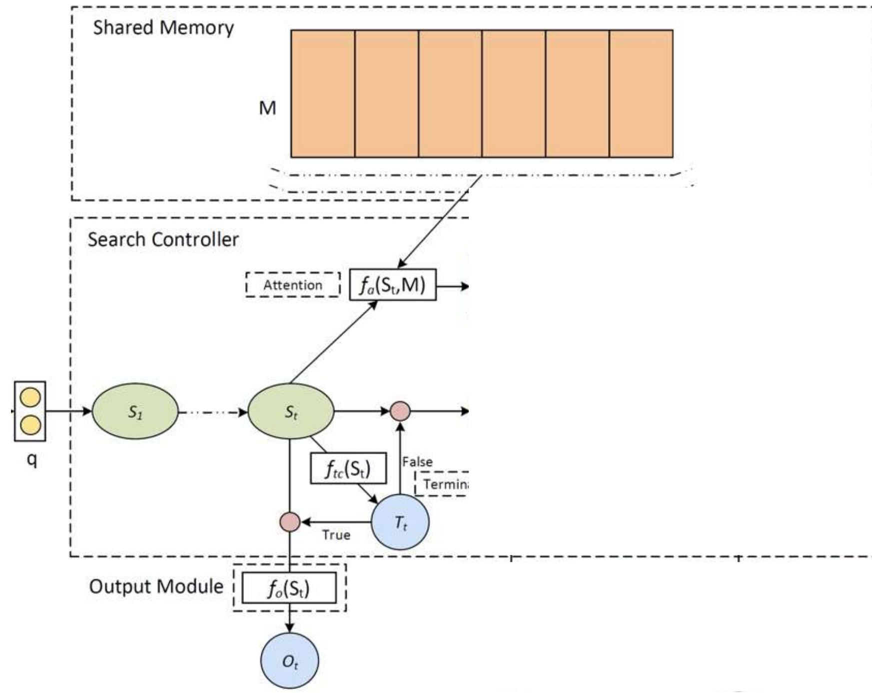
[Shen+ 17]

ReasoNet: learn to stop reading

Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and Von Miller for Denver.
Answer	**Von Miller**

Rank-1
Rank-2
Rank-3

S: Who was the #2 pick in the 2011 NFL Draft?



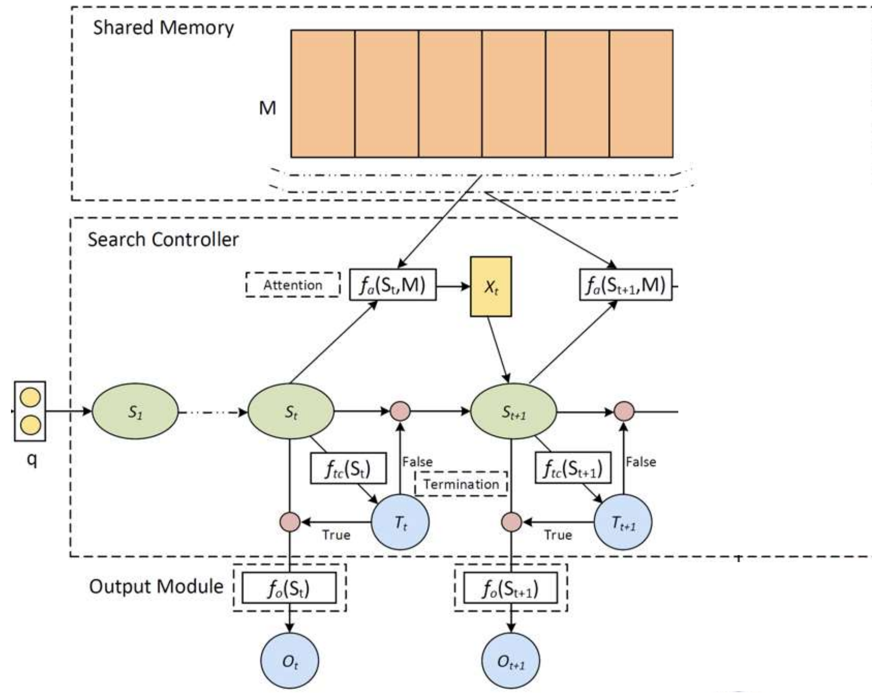
Step	Termination Probability	Prob. Answer
1	0.001	0.392

ReasoNet: learn to stop reading

Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and Von Miller for Denver.
Answer	**Von Miller**

Rank-1
Rank-2
Rank-3

S: Manning is #1 pick of 1998, but this is unlikely the answer.



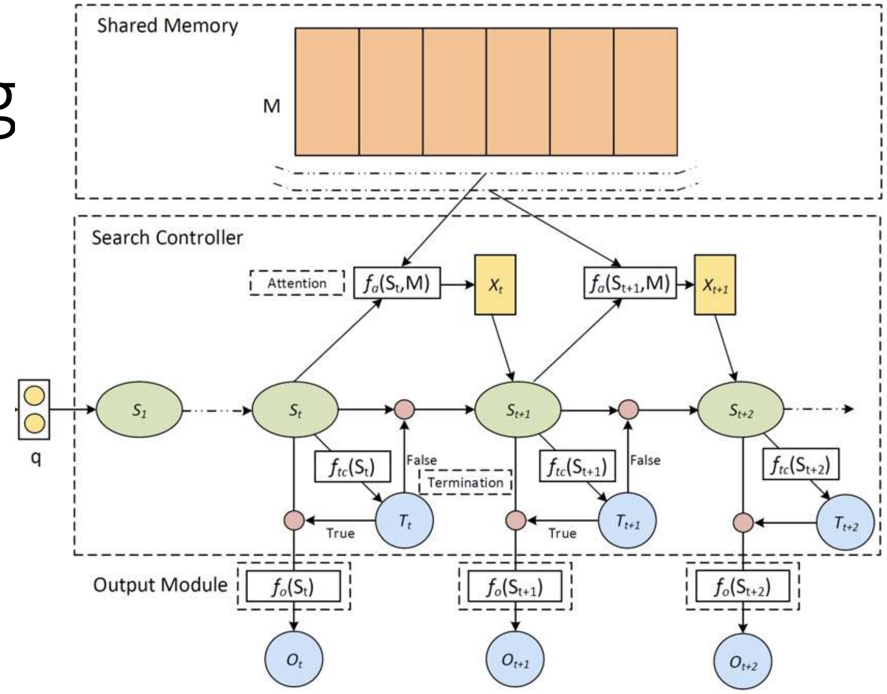
Step	Termination Probability	Prob. Answer
1	0.001	0.392
2	0.675	0.649

ReasoNet: learn to stop reading

Query	Who was the #2 pick in the 2011 NFL Draft?
Passage	Manning was the #1 selection of the 1998 NFL draft, while Newton was picked first in 2011. The matchup also pits the top two picks of the 2011 draft against each other: Newton for Carolina and Von Miller for Denver.
Answer	**Von Miller**

Rank-1
Rank-2
Rank-3

S: Manning is #1 pick of 1998, Newton is #1 pick of 2011, but neither is the answer.



Step t	Termination Probability f_{tc}	Prob. Answer f_o
1	0.001	0.392
2	0.675	0.649
3	0.939	0.865

Stochastic Answer Net

- Training uses *stochastic prediction dropout* on the answer module
- Reasoning employs all the outputs of multiple step reasoning via voting
- Best documented MRC model on the SQuAD leaderboard as of Dec. 19, 2017

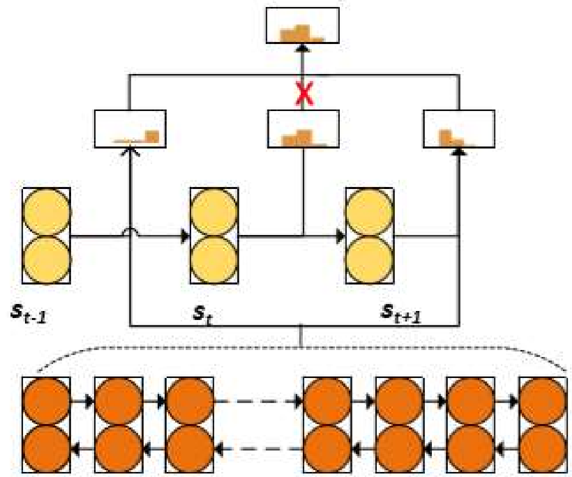


Figure 1: Illustration of "stochastic prediction dropout" in the answer module during training. At each reasoning step t, the model combines memory (bottom row) with hidden states s_{t-1} to generate a prediction (multinomial distribution). Here, there are three steps and three predictions, but one prediction is dropped and the final result is an average of the remaining distributions.

[Liu+ 18]

Interim summary

- Knowledge base QA tasks
  - From semantic parsing to embedding-based approaches
  - Reasoning on symbolic or neural space
  - Multi-turn knowledge base QA agents
- Neural MRC models for text-based QA
  - MRC tasks, e.g., SQuAD, MS MARCO
  - Three embedding components: word/context/query-context
  - Multi-step reasoning

Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogue**
  - **Overview**
  - Traditional approaches
  - Natural language understanding and dialogue state tracking
  - Deep RL for dialogue policy learning
  - Neural approaches and E2E learning
- Part 4: Fully data-driven conversation models and chatbots

An Example Dialogue with Movie-Bot

```
Turn 0 usr: can i get 2 tickets for race
Turn 1 sys: What date would you like to watch it?
Turn 2 usr: tomorrow
Turn 3 sys: Which theater would you like?
Turn 4 usr: amc pacific place 11 theater
Turn 5 sys: Which city you would like?
Turn 6 usr: seattle
Turn 7 sys: What time would you like to see it?
2 tickets for you
cific place 11 theater
```

Actual dialogues can be more complex:

- Natural language understanding errors
  - Input may be spoken language form
  - Need to reason under uncertainty
- Constraint violation
  - Revise information collected earlier
- …

iuLab/TC-Bot

Slot-Filling Dialogues

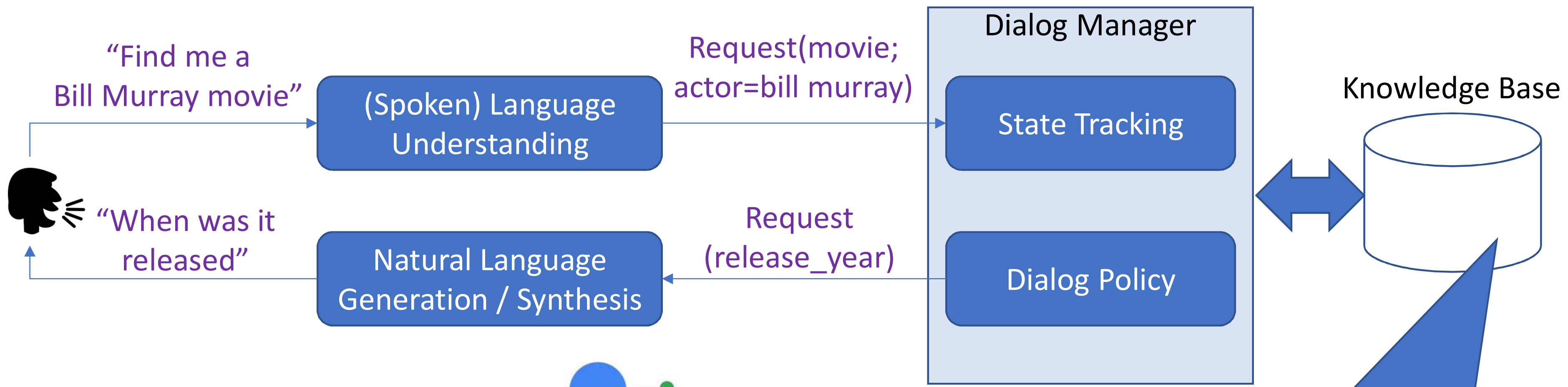
- **Domain**: movie, restaurant, flight, …
- **Slot**: information to be filled in before completing a task
  - For Movie-Bot: movie-name, theater, number-of-tickets, price, …
- **Intent** (dialogue act):
  - Inspired by speech act theory (communication as action)
    request, confirm, inform, thank-you, …
  - Some may take parameters:
    thank-you(), request(price), inform(price=$10)

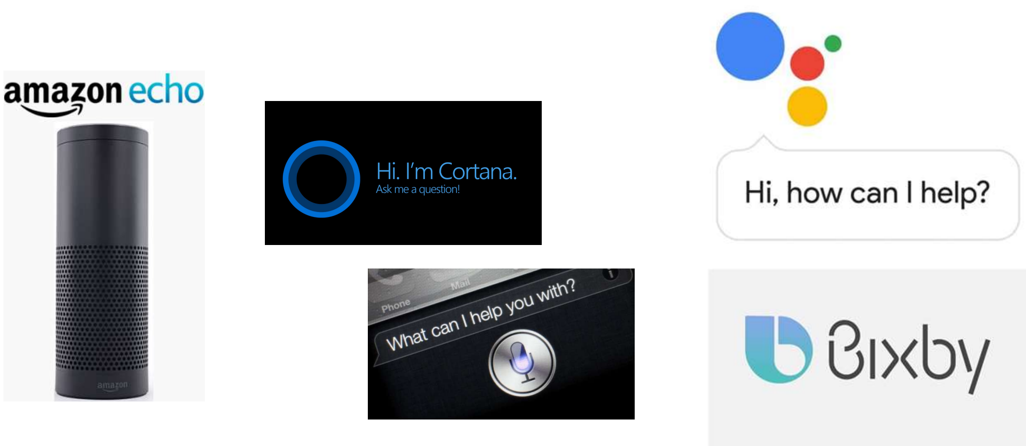
"Is Kungfu Panda the movie you are looking for?"

⇩

confirm(moviename="kungfu panda")

A Multi-turn Task-oriented Dialogue Architecture

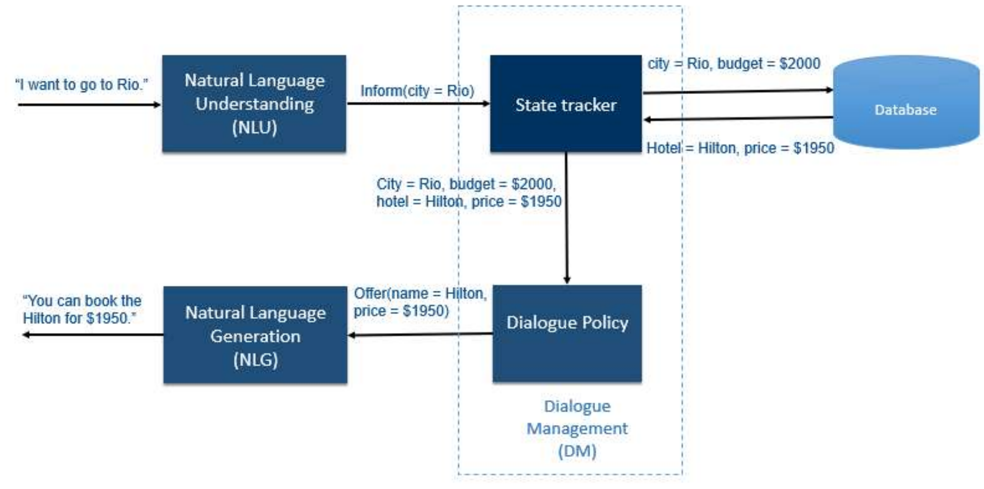


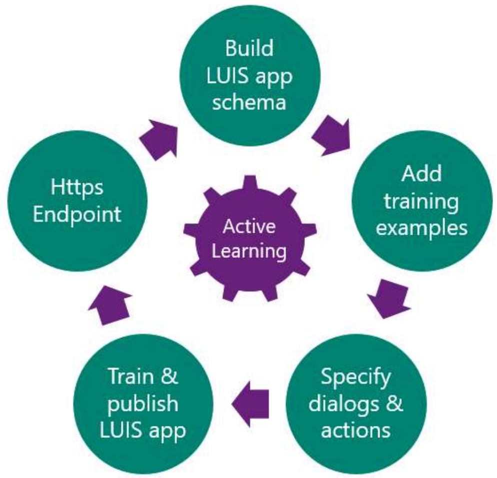


Entity-Centric Knowledge Base

Movie	Actor	Release Year
Groundhog Day	Bill Murray	1993
Australia	Nicole Kidman	X
Mad Max: Fury Road	X	2015

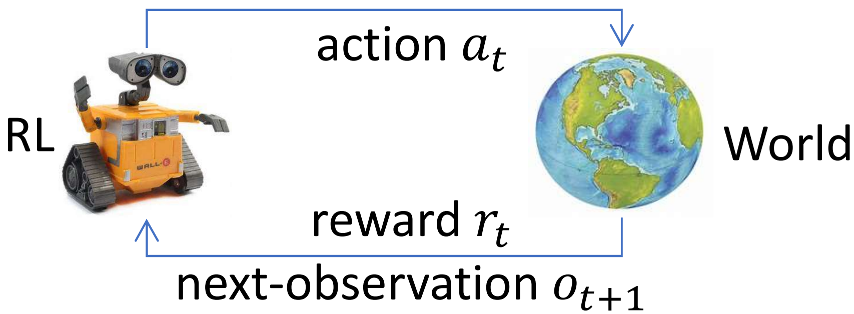
Task-oriented Dialogue Agents via LUIS/TCP/BF





- LUIS -> NLU, trained on labeled data, not robust to paragraphing
- TCP -> DM, hand-crafted, not flexible and cannot handle corner cases
- (BF -> multiple channels)
- **Active research: Improving agents E2E via Reinforcement Learning (RL)**

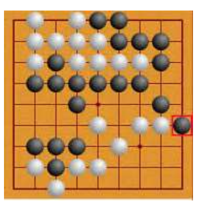
Reinforcement Learning (RL)

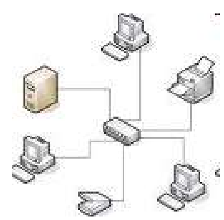


Typical goal of RL

At each step t, given history so far, take action a_t to maximize long-term reward ("return"):

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \cdots$$





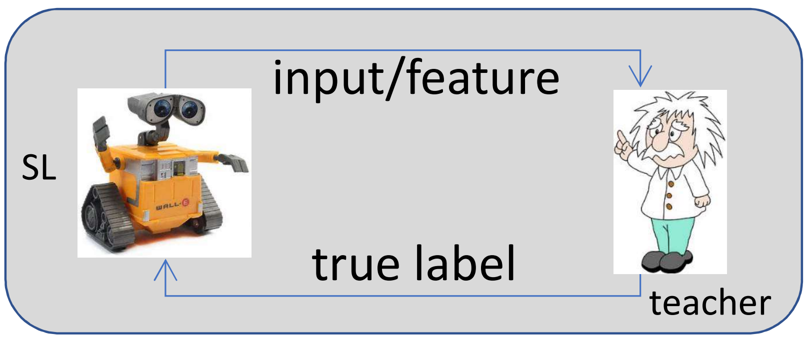
......

"Reinforcement Learning: An Introduction", 2nd ed., Sutton & Barto

RL vs. SL (supervised learning)

Differences from supervised learning

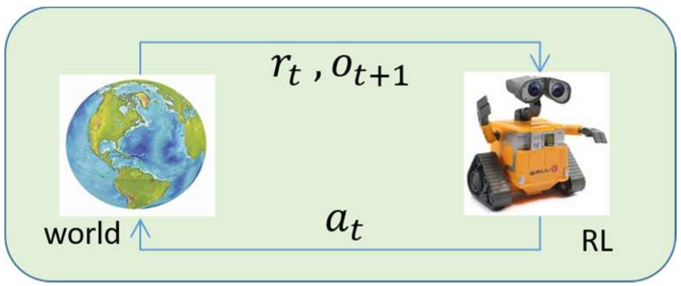
- Learn by trial-and-error ("experimenting")
  - Need efficient exploration
- Optimize long-term reward ($r_1 + \gamma r_2 + \cdots$)
  - Need temporal credit assignment

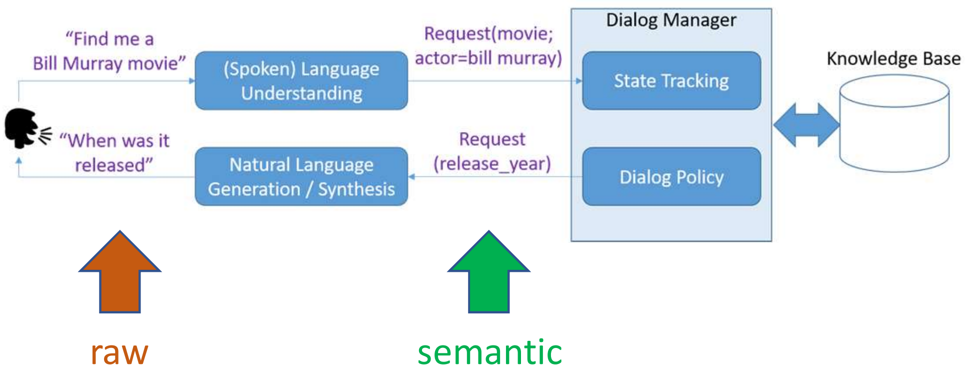


Similarities to supervised learning

- Generalization and representation
- Hierarchical problem solving
- …

Conversation as RL





- Observation and action
  - Raw representation (utterances in natural language form)
  - Semantic representation (intent-slot-value form)
- Reward
  - +10 upon successful termination
  - -10 upon unsuccessful termination
  - -1 per turn
  - …

Pioneered by [Levin+ 00]
Other early examples: [Singh+ 02; Pietquin+ 04; Williams&Young 07; etc.]

Dialogue System Evaluation

- **Metrics**: what numbers matter?
  - Success rate: #Successful_Dialogues / #All_Dialogues
  - Average turns: average number of turns in a dialogue
  - User satisfaction
  - Consistency, diversity, engaging, ...
  - Latency, backend retrieval cost, …

- **Methodology**: how to measure those numbers?

Methodology (I): Human Subject Study

- Recruit human subjects, randomly split into two groups (A and B)
- Blindly test two systems on these groups
- Compare metrics of interests between A and B

- Pros
  - Real humans involved
- Cons
  - Expensive and time-consuming, does not scale well
  - Lab users and actual users may behave differently
  - Limited use, mostly for comparing two (or few) given systems

Methodology (II): Actual users

- Deploy the dialogue system to serve actual users
- Keep track of metrics of interest from user logs

- Pros
  - Arguably the gold standard evaluation setup
- Cons
  - Mostly feasible for major companies with a large user base
  - Expensive and risky (potential business disruptions)

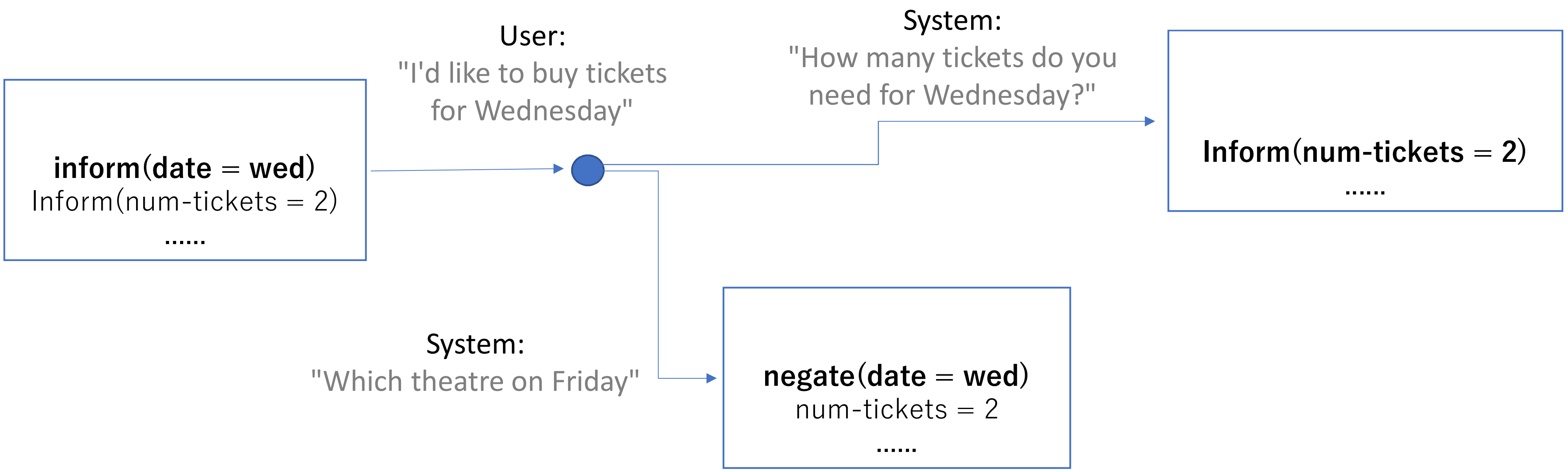
Methodology (III): Simulated users

- Create a simulated user U: conversational history -> next utterance
  - Agenda based [Schatzmann & Young 09]
  - Model based [El Asri+ 16]
- Can run any policy or RL algorithm against U to measure metrics

- Pros
  - Cheap to run, with *unlimited* amount of data
- Cons
  - User simulation is only a (usually rough) approximation of real users
  - Building a good user simulator is nontrivial

Agenda-based Simulated User [Schatzmann & Young 09]

- User state consists of (agenda, goal); goal is fixed throughout dialogue
- Agenda is maintained (stochastically) by a first-in-last-out stack



Implementation of a simplified user simulator: https://github.com/MiuLab/TC-Bot

Methodology: Summary

	Lab user subjects	Actual users	Simulated users
Truthfulness		✓	✗
Scalability	✗	✓	✓
Flexibility	✗		✓
Expense	✗		✓
Risk	✓	✗	✓

A Hybrid Approach

User Simulation

⇩

Small-scale Human Evaluation
(lab, Mechanical Turk, …)

⇩

Large-scale Deployment
(optionally with continuing
incremental refinement)

Outline

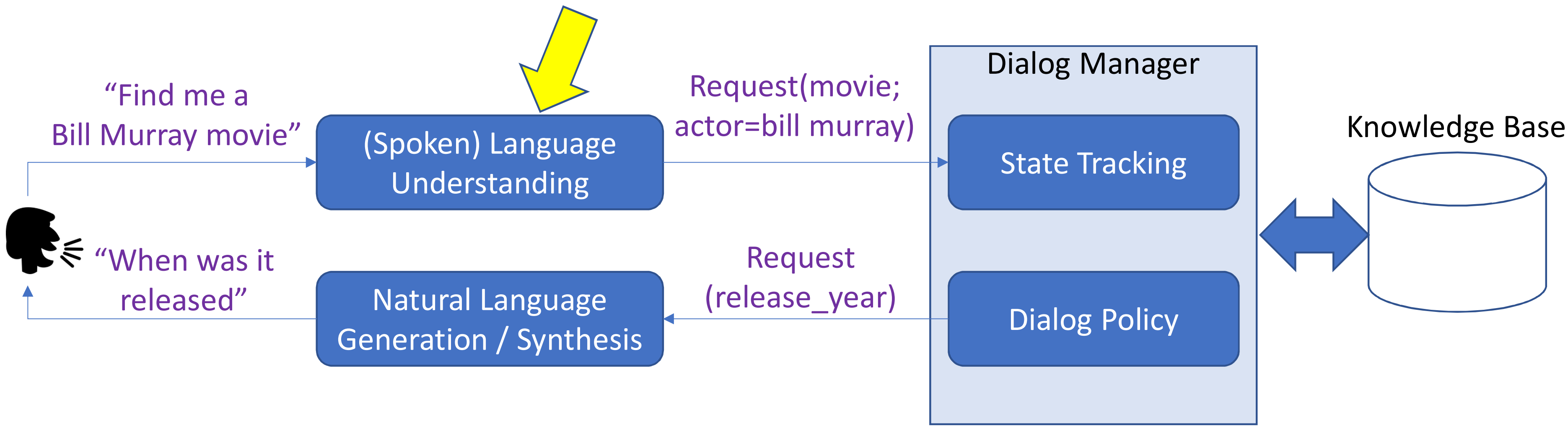
- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogue**
    - Overview
    - **Traditional approaches**
    - Natural language understanding and dialogue state tracking
    - Deep RL for dialogue policy learning
    - Neural approaches and E2E learning
- Part 4: Fully data-driven conversation models and chatbots

Decision-theoretic View of Dialogue Management

- Principle of Maximum Expected Utility (MEU):

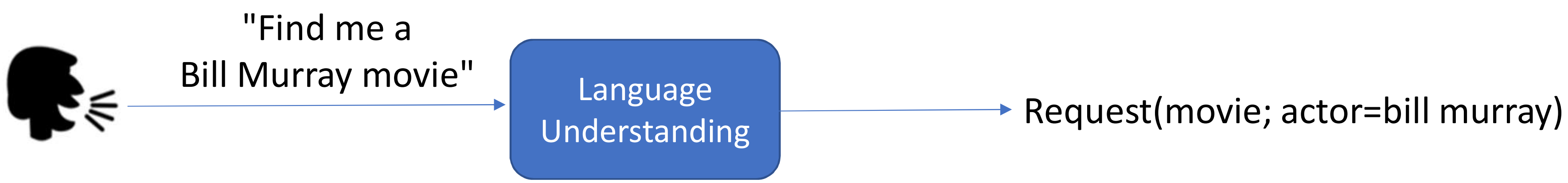
  An optimal decision maximizes the expected utility of outcomes.

- Same in RL that typically aims to maximize expected total reward



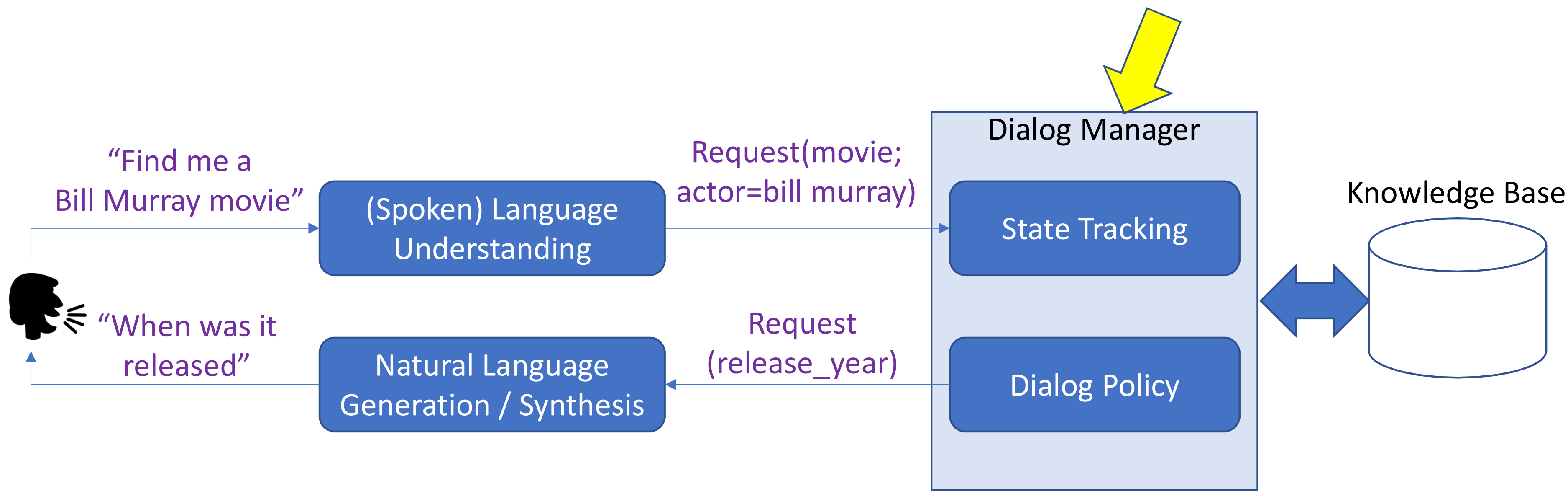
Language Understanding in Dialogue Systems

- Goal is to predict (domain, intent, slot) from user utterance



- Often treated as a classification problem
- Solved by SVM, CRF, generative models (e.g., HMM), ...,
- See surveys by Wang et al. [2005] and De Mori et al. [2008]
- More to cover in the next section

Decision-theoretic View of Dialogue Management



Dialogue Management

- Typical assumptions in earlier work: small number of hand-coded states

Feature	Values	Explanation
Greet (G)	0,1	Whether the system has greeted the user
Attribute (A)	1,2,3,4	Which attribute is being worked on
Confidence/Confirmed (C)	0,1,2,3,4	0,1,2 for low, medium, and high ASR confidence. 3,4 for explicitly confirmed, and disconfirmed
Value (V)	0,1	Whether value has been obtained for current attribute
Tries (T)	0,1,2	How many times current attribute has been asked
Grammar (M)	0,1	Whether non-restrictive or restrictive grammar was used
History (H)	0,1	Whether there was trouble on any previous attribute

(From NJFun [Singh+ 00])

- Q(s,a) measures the highest long-term reward, updated by

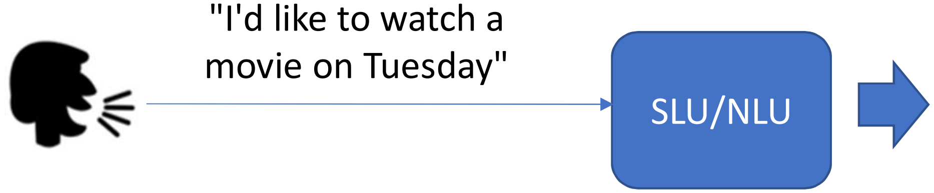
$$Q(s,a) \leftarrow Q(s,a) + \eta \left(r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right)$$

- Once learned, action in state s is $\arg\max_{a} Q(s,a)$

[Levin+ 00; Walker+ 98; Singh+ 00; Scheffler & Young 02; etc.]

Language Understanding Uncertainty

- True state of a conversation is partially observable



Confidence	Result
0.9	Inform(date="Tuesday")
0.1	Inform(date="Thursday")
0.05	Inform(moviename="Tuesday")
...	...

"N-best list"

- Partially Observable MDP (POMDP) as a principled framework
  Key idea: maintains a posterior distribution of current dialogue state
  [Roy+ 00; Zhang+ 01; Williams&Young 07, etc.]
- Limitations: (1) scalability, (2) domain knowledge required

Scaling up Dialogue Optimization

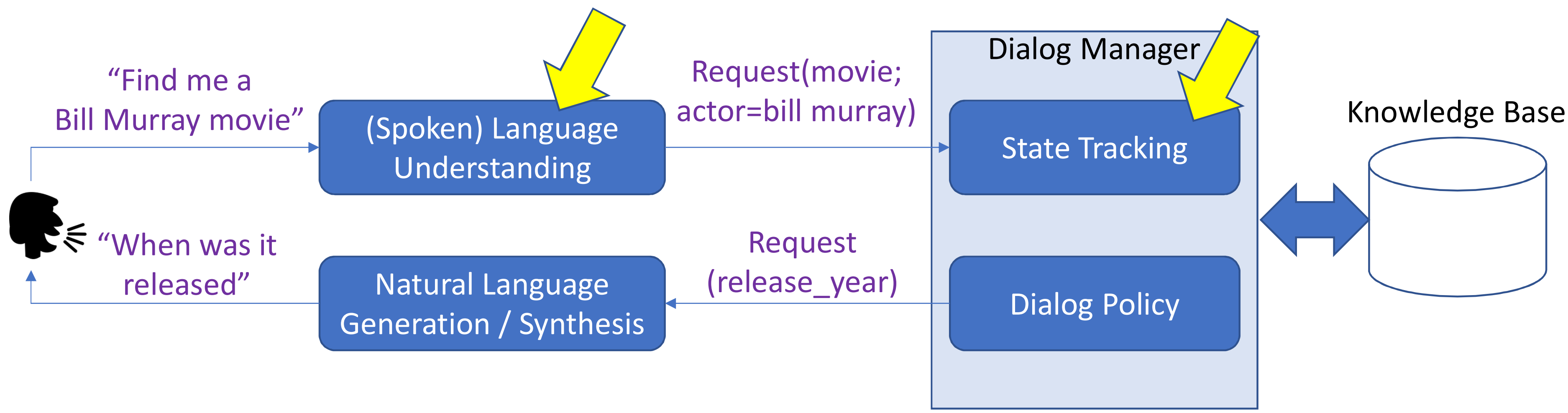
- Use approximate POMDP algorithms leveraging problem-specific structures
    - Augmented MDP [Roy et al., 2000]
    - Summary POMDPs [Williams&Young 07]

- Use RL algorithms with function approximation
    - Requires human designed features that implicitly encode domain knowledge
    - Use of linear Q-functions [Henderson et al., 2008]
    - Automated feature selection for improved scalability [Li et al., 2009]
    - Nonparametric Gaussian processes [Gasic+ 14]

Outline

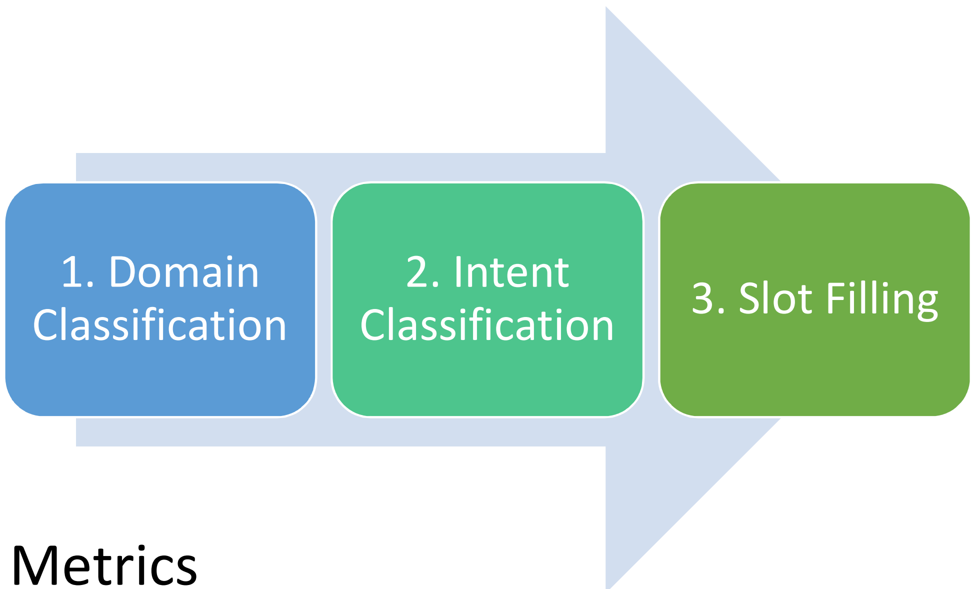
- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogue**
  - Overview
  - Traditional approaches
  - **Natural language understanding and dialogue state tracking**
  - Deep RL for dialogue policy learning
  - Neural approaches and E2E learning
- Part 4: Fully data-driven conversation models and chatbots

Decision-theoretic View of Dialogue Management



Language Understanding

- Often a multi-stage pipeline



W	find	recent	comedies	by	james	cameron
	↓	↓	↓	↓	↓	↓
S	O	B-date	B-genre	O	B-dir	I-dir
D	movies					
I	find_movie					

Figure 1: *An example utterance with annotations of semantic slots in IOB format (S), domain (D), and intent (I), B-dir and I-dir denote the director name.*

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

LU – Domain/Intent Classification

Mainly viewed as an utterance classification task

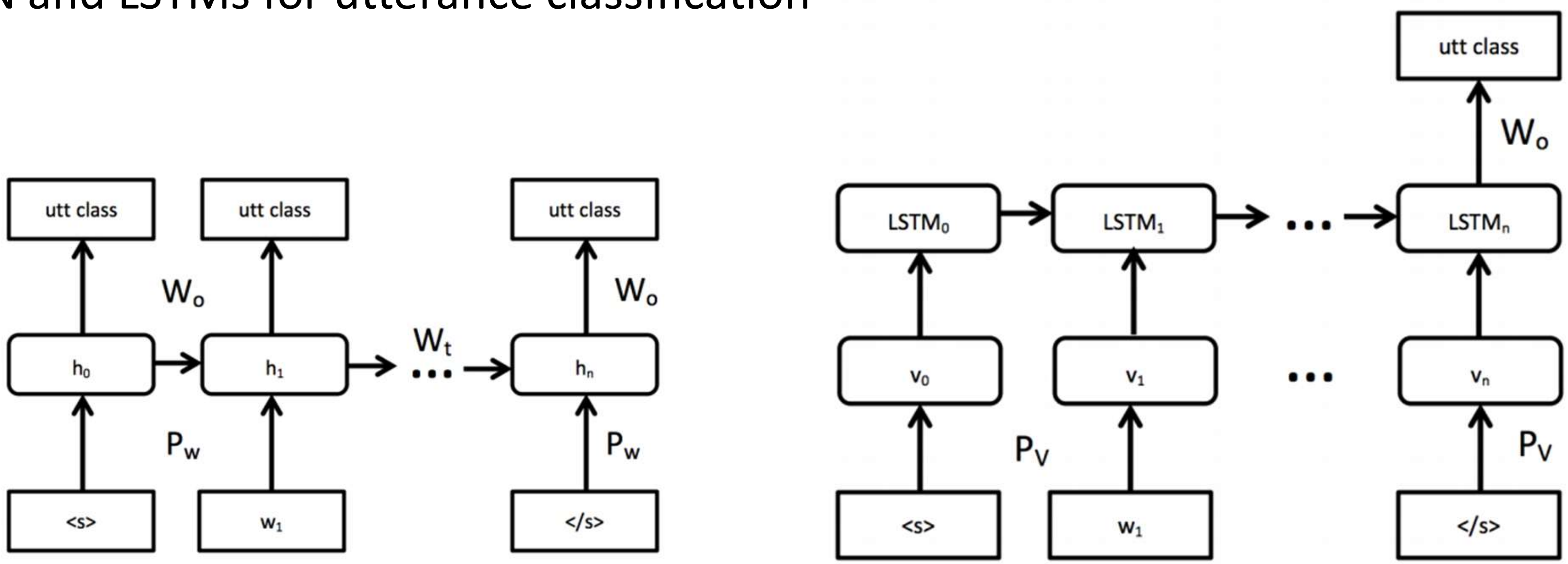
- Given a collection of utterances u_i with labels c_i, D= $\{(u_1,c_1),\ldots,(u_n,c_n)\}$ where $c_i \in C$, train a model to estimate labels for new utterances u_k.

find me a cheap taiwanese restaurant in oakland

Movies	Find_movie
Restaurants	Buy_tickets
Sports	Find_restaurant
Weather	Book_table
Music	Find_lyrics
…	…

DNN for Domain/Intent Classification (Ravuri & Stolcke, 2015)

- RNN and LSTMs for utterance classification



Intent decision after reading all words performs better

LU – Slot Filling

As a sequence tagging task

- Given a collection tagged word sequences, $S=\{((w_{1,1},w_{1,2},\ldots, w_{1,n1}), (t_{1,1},t_{1,2},\ldots,t_{1,n1})), ((w_{2,1},w_{2,2},\ldots,w_{2,n2}), (t_{2,1},t_{2,2},\ldots,t_{2,n2})) \ldots\}$ where $t_i \in M$, the goal is to estimate tags for a new word sequence.

flights from Boston to New York today

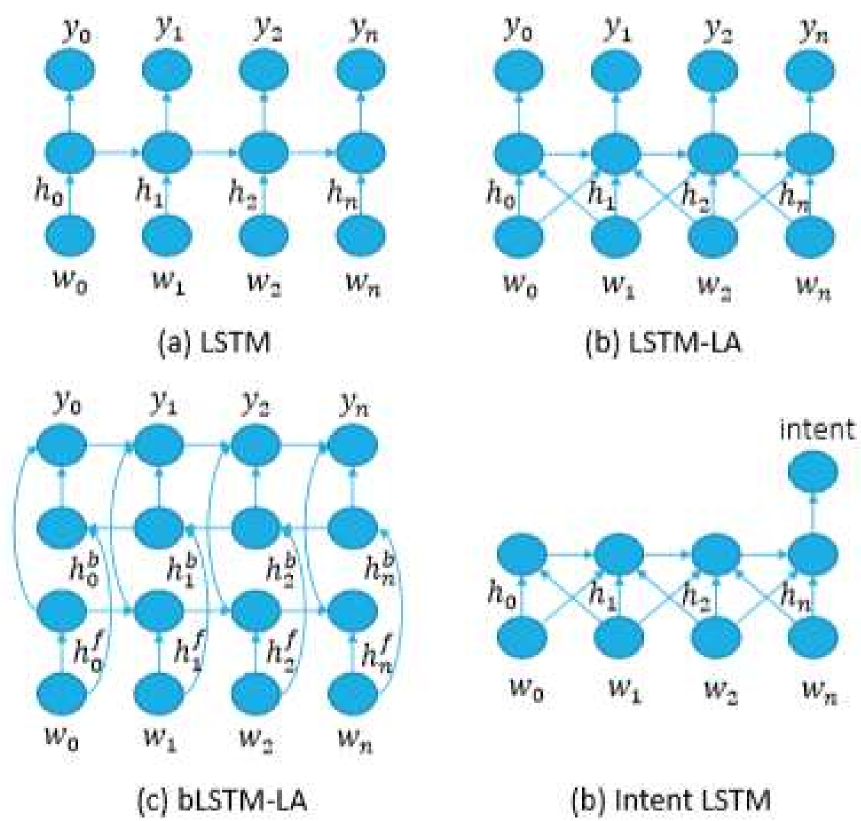
	flights	from	Boston	to	New	York	today
Entity Tag	O	O	B-city	O	B-city	I-city	O
Slot Tag	O	O	B-dept	O	B-arrival	I-arrival	B-date

RNN for Slot Tagging – I [Hakkani-Tur+ 16]

- Variations:
  a. RNNs with LSTM cells
  b. Look-around LSTM
  c. Bi-directional LSTMs
  d. *Intent LSTM*

- May also take advantage of …
  - whole-sentence information
  - multi-task learning
  - contextual information

- For further details on NLU, see this IJCNLP tutorial by Chen & Gao.



E2E MemNN for Contextual LU [Chen+ 16]

1. Sentence Encoding

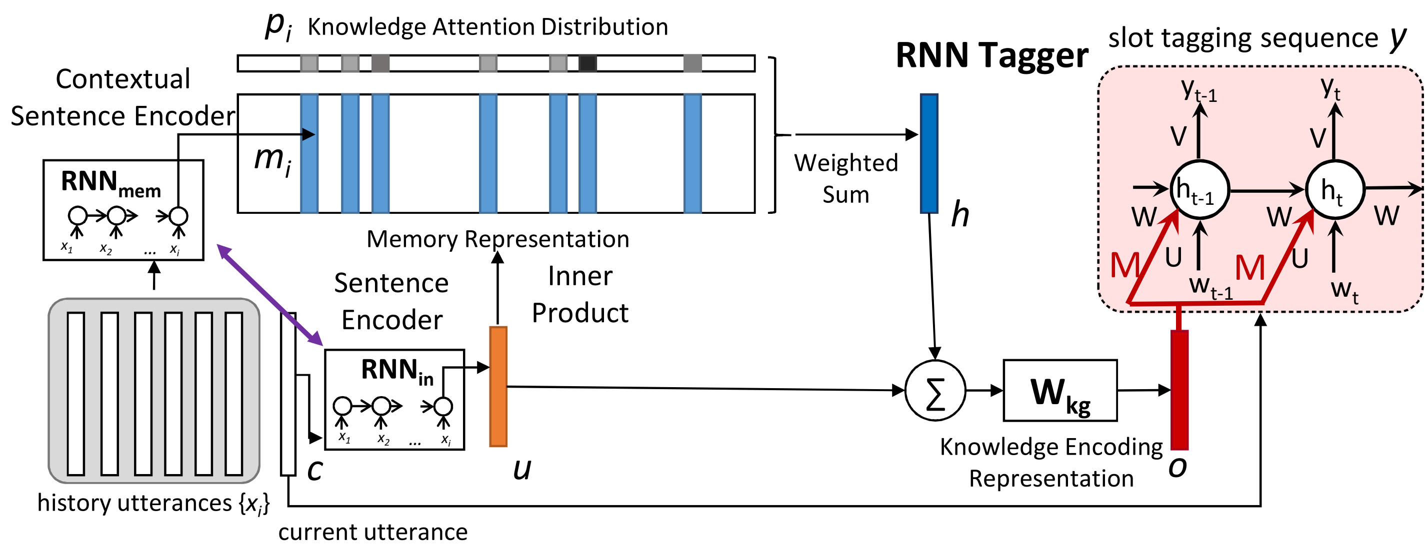
$$m_i = \mathrm{RNN}_{\mathrm{mem}}(x_i)$$
$$u = \mathrm{RNN}_{\mathrm{in}}(c)$$

2. Knowledge Attention

$$p_i = \mathrm{softmax}(u^T m_i)$$

3. Knowledge Encoding

$$h = \sum_i p_i m_i \quad o = W_{\mathrm{kg}}(h + u)$$

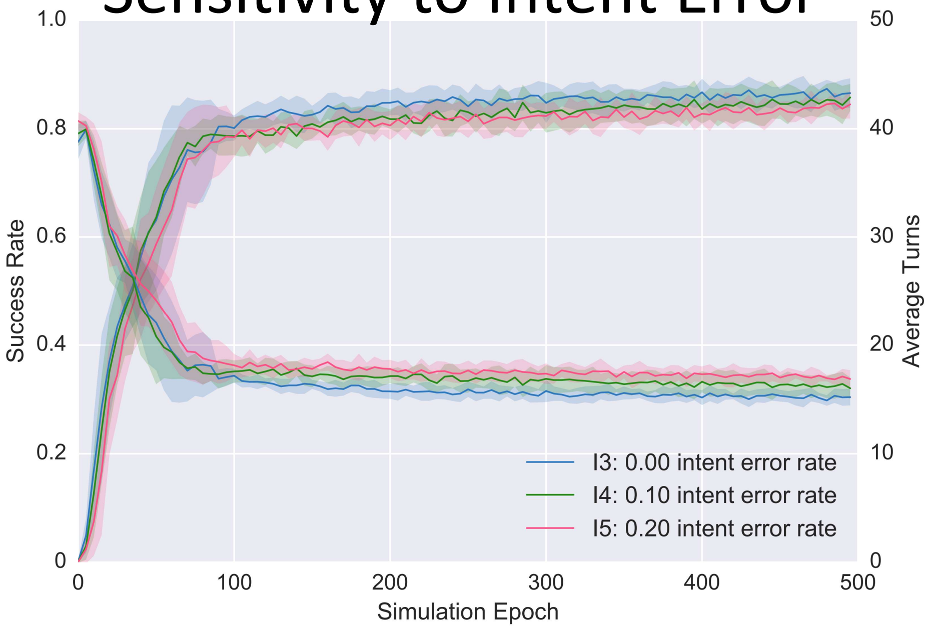


Idea: additionally incorporating contextual knowledge during slot tagging
→ track dialogue states in a latent way

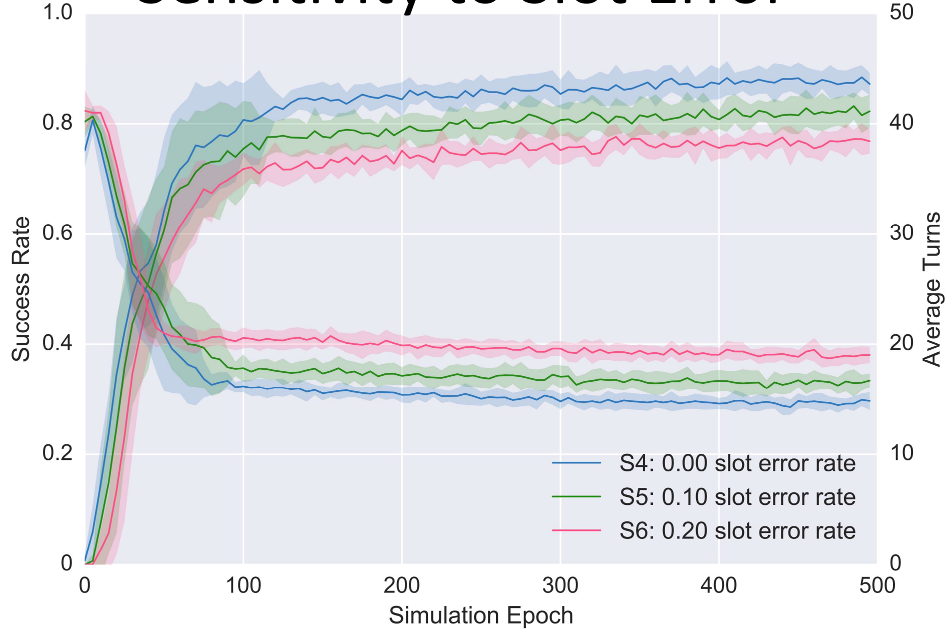
LU Importance [Li+ 17]

- Compare different types of LU errors

Sensitivity to Intent Error



Sensitivity to Slot Error



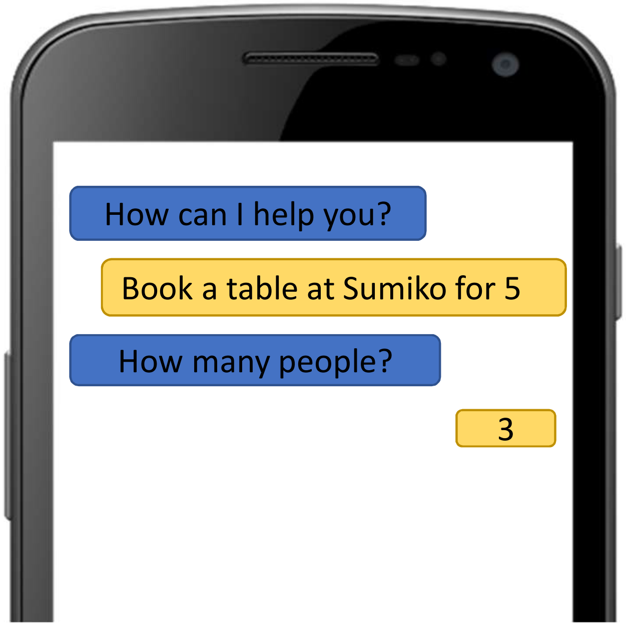
Slot filling is more important than intent detection in language understanding

Dialogue State Tracking (DST)

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

Slot	Value
# people	5 (0.5)
time	5 (0.5)

Slot	Value
# people	3 (0.8)
time	5 (0.8)



Multi-Domain Dialogue State Tracking (DST)

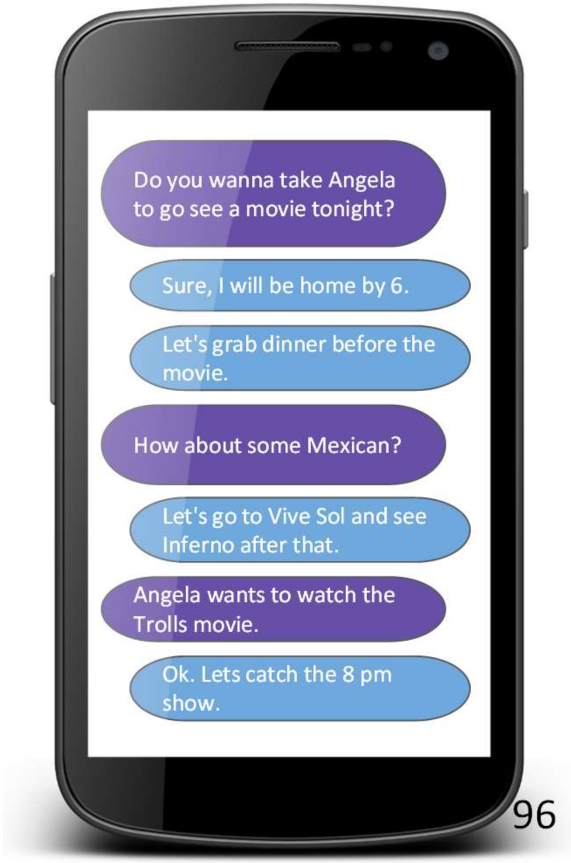
- A full representation of the system's belief of the user's goal at any point during the dialogue
- Used for making API calls

Movies

11/15/16			
6 pm	7 pm	8 pm	9 pm
2	3		
Inferno	Trolls		
Century 16			

Restaurants

Date	11/15/16		
Time	6:30 pm	7 pm	7:30 pm
Cuisine	Mexican		
Restaurant	Vive Sol		



Dialog State Tracking Challenge (DSTC)

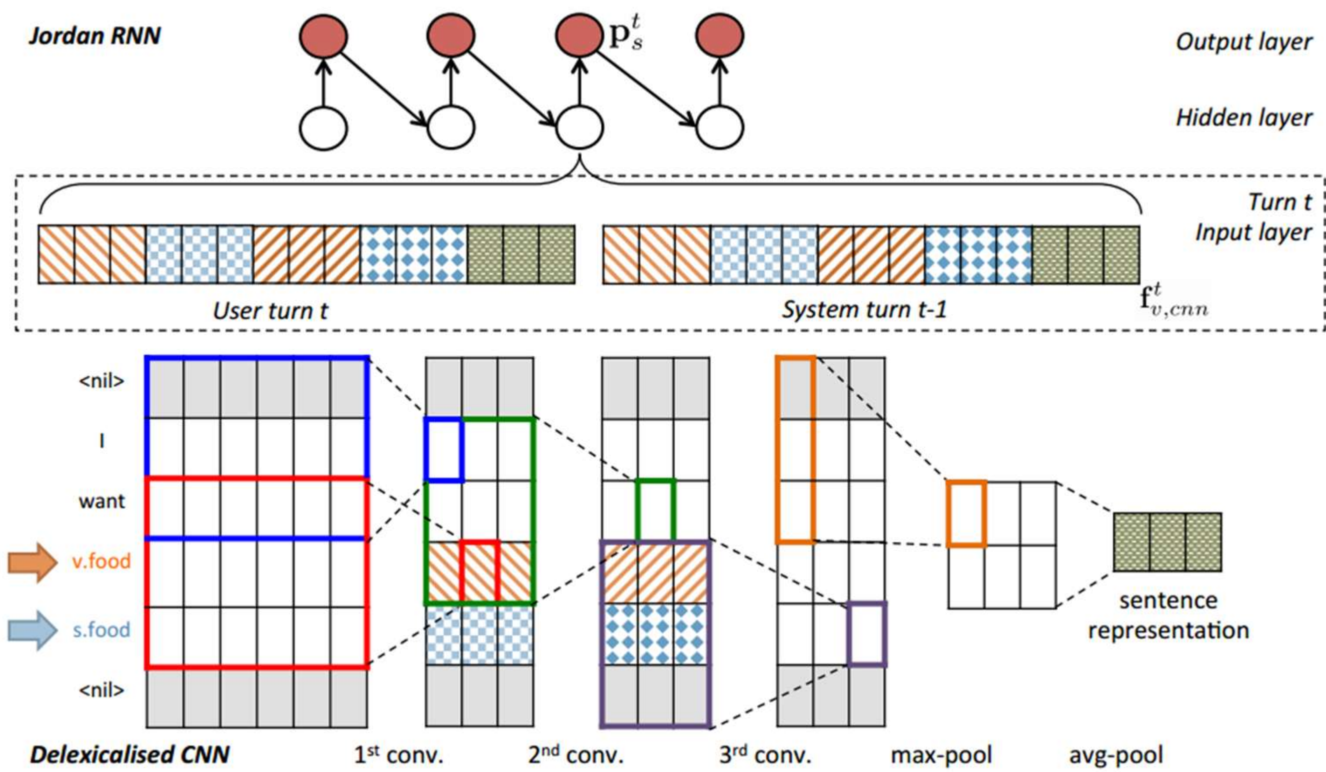
(Williams et al. 2013, Henderson et al. 2014, Henderson et al. 2014, Kim et al. 2016, Kim et al. 2016)

Challenge	Type	Domain	Data Provider	Main Theme
DSTC1	Human-Machine	Bus Route	CMU	Evaluation Metrics
DSTC2	Human-Machine	Restaurant	U. Cambridge	User Goal Changes
DSTC3	Human-Machine	Tourist Information	U. Cambridge	Domain Adaptation
DSTC4	Human-Human	Tourist Information	I2R	Human Conversation
DSTC5	Human-Human	Tourist Information	I2R	Language Adaptation

DST Evaluation

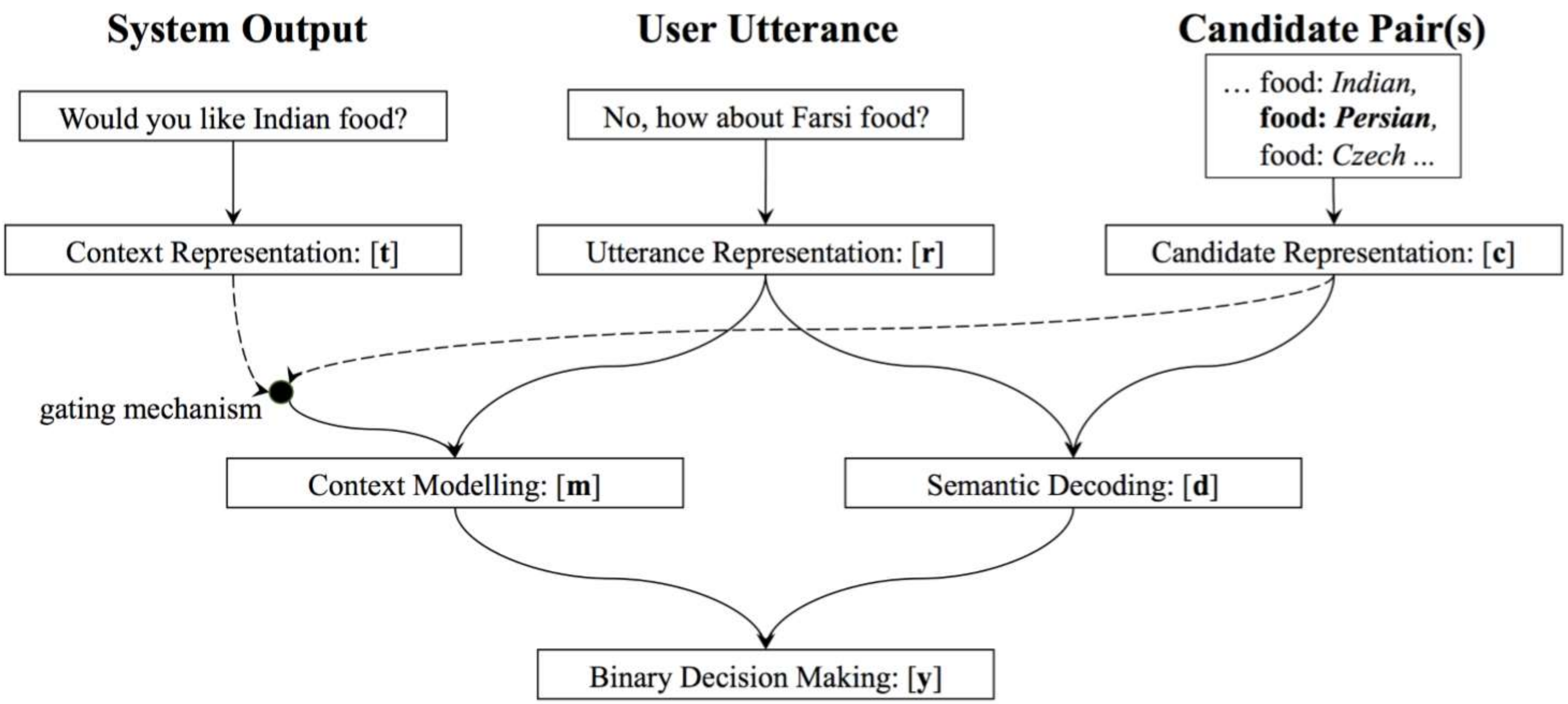
- Dialogue State Tracking Challenges
  - DSTC2-3, human-machine
  - DSTC4-5, human-human
- Metric
  - Tracked state accuracy with respect to user goal
  - Recall/Precision/F-measure individual slots

NN-Based DST [Henderson+ 13; Mrkšić+ 15; Mrkšić+ 16]



(Figure from Wen+ 17)

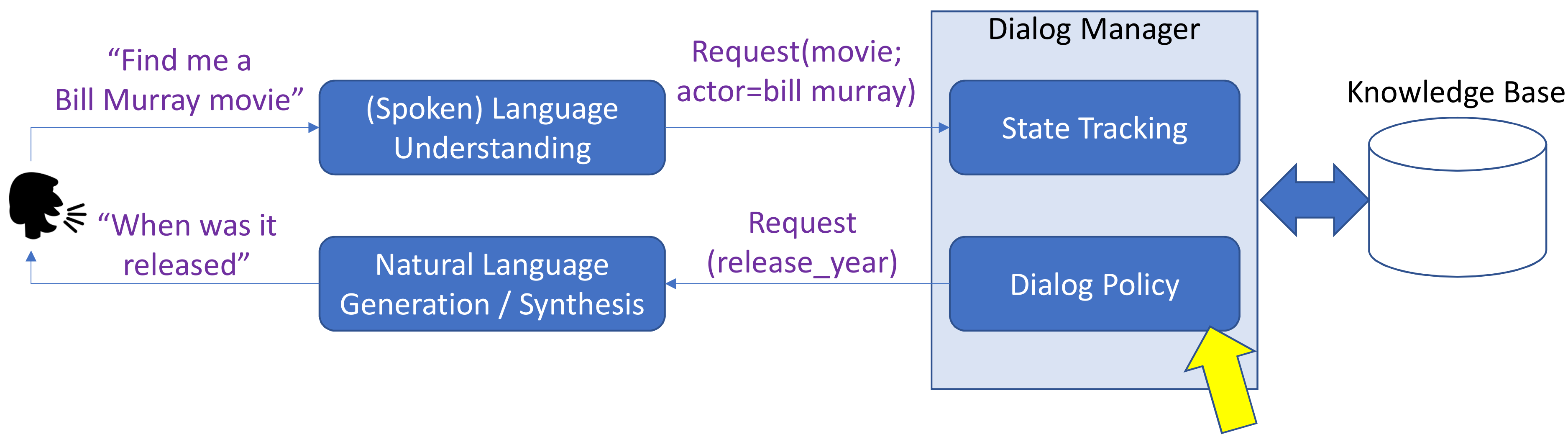
Neural Belief Tracker [Mrkšić+ 16]



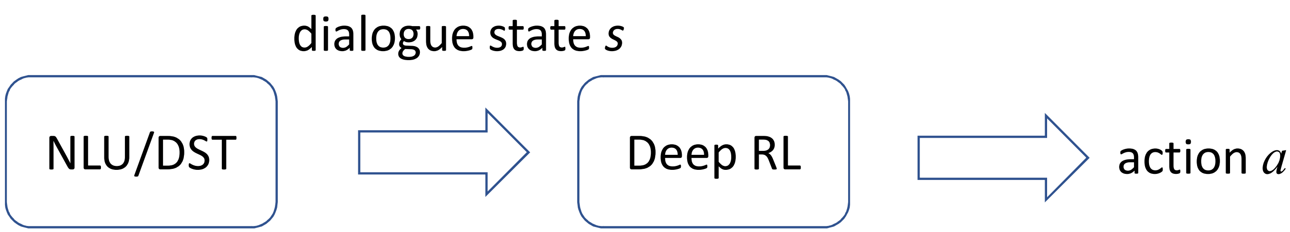
Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogue**
  - Overview
  - Traditional approaches
  - Natural language understanding and dialogue state tracking
  - **Deep RL for dialogue policy learning**
  - Neural approaches and E2E learning
- Part 4: Fully data-driven conversation models and chatbots

Decision-theoretic View of Dialogue Management

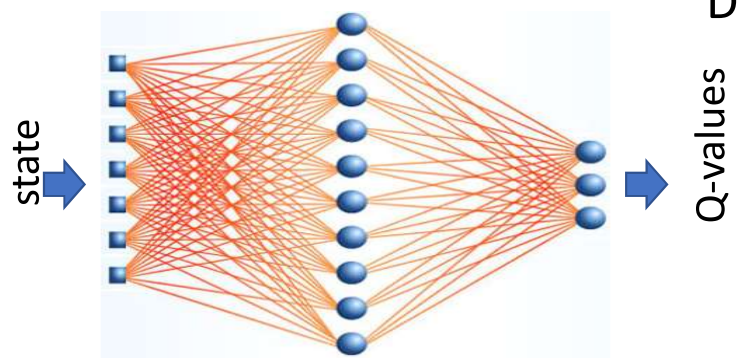


Deep RL for Dialogue Policy Optimization



- Action may be semantic (confirm(date="tuesday"))
  or in NLP form ("is it for tuesday?")
  or a backend operation (e.g., API call, KB lookup)
- Two main classes of RL algorithms
  - Value function based: Q-learning (DQN), Sarsa, ...
  - Policy based: REINFORCE, policy gradient / actor-critic, ...

Policy Optimization with DQN



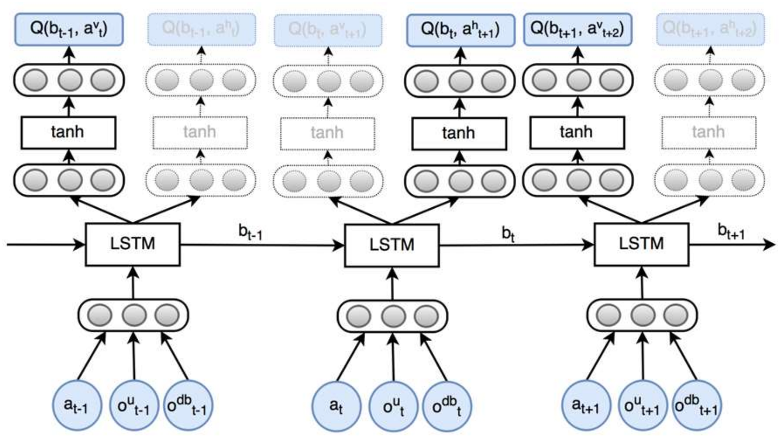
DQN-learning of network weights θ: apply SGD to solve

$$\hat{\theta} \leftarrow \arg\min_{\theta} \sum_{t} \left(r_{t+1} + \gamma \max_{a} Q_T(s_{t+1}, a) - Q_L(s_t, a_t) \right)^2$$

"Target network" to synthesize regression target

"Learning network" whose weights are to be updated

[Mnih+ 15]



RNN/LSTM may be used to implicitly track states (without a separate dialogue state tracker)
[Zhao & Eskenazi 16]

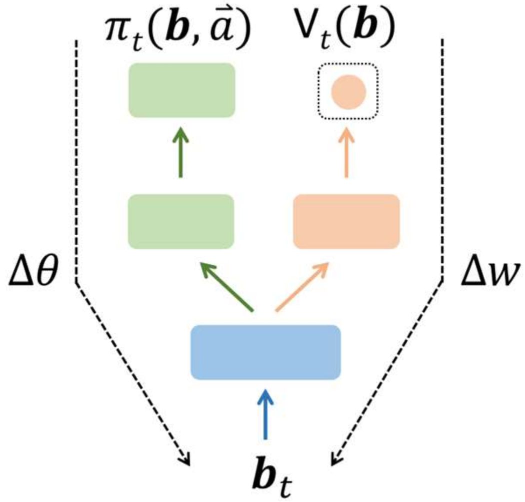
Policy Optimization with Policy Gradient

- PG does gradient descent in policy parameter space to improve reward

$$\nabla_\theta J(\theta) = \mathbb{E}\left[\nabla_\theta \log \pi_\theta(a|\mathbf{b}) Q^{\pi_\theta}(\mathbf{b}, a)\right]$$

- REINFORCE [Williams 1992]: simplest PG algorithm

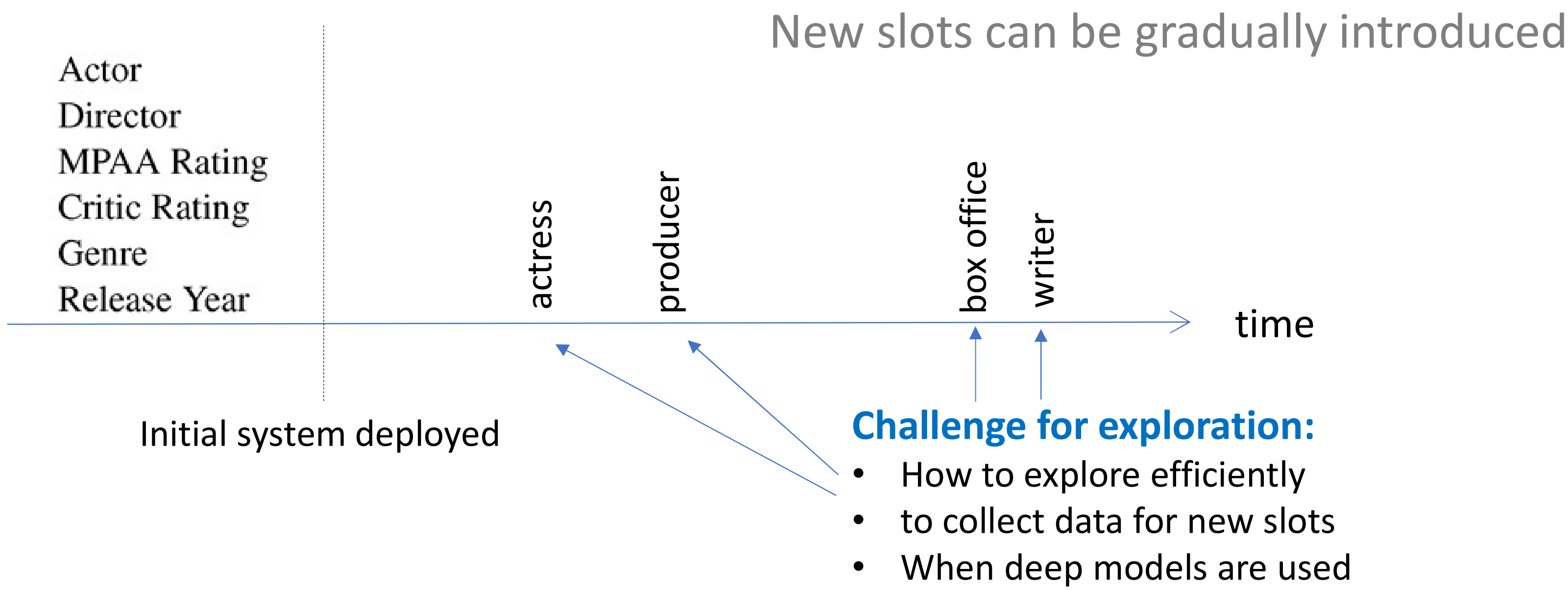
- Advantaged Actor-Critic (A2C) / TRACER
  - w: updated by least-squared regression
  - θ: updated as in PG



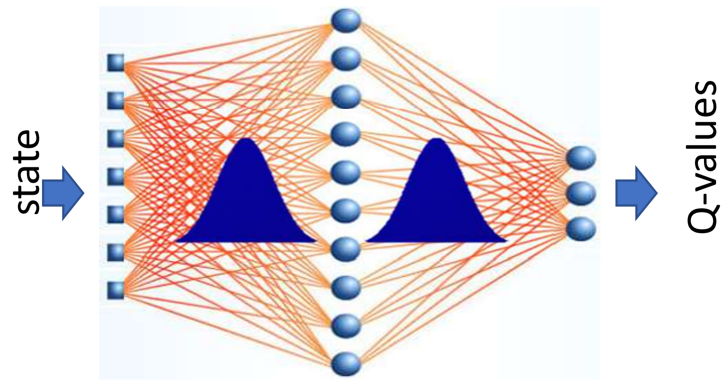
A2C/TRACER [Su+ 17]

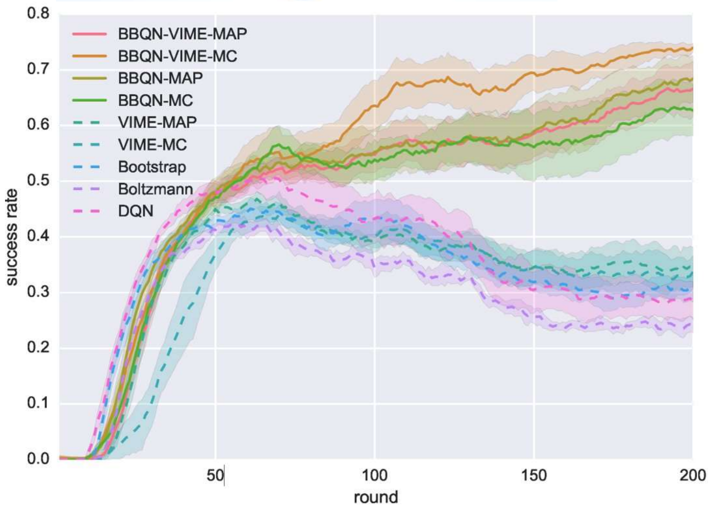
Domain Extension and Exploration

- Most goal-oriented dialogs require a closed and well-defined domain
- Hard to include all domain-specific information up-front



Bayes-by-Backprop Q (BBQ) network





BBQ-learning of network params $\theta = (\mu, \sigma^2)$:

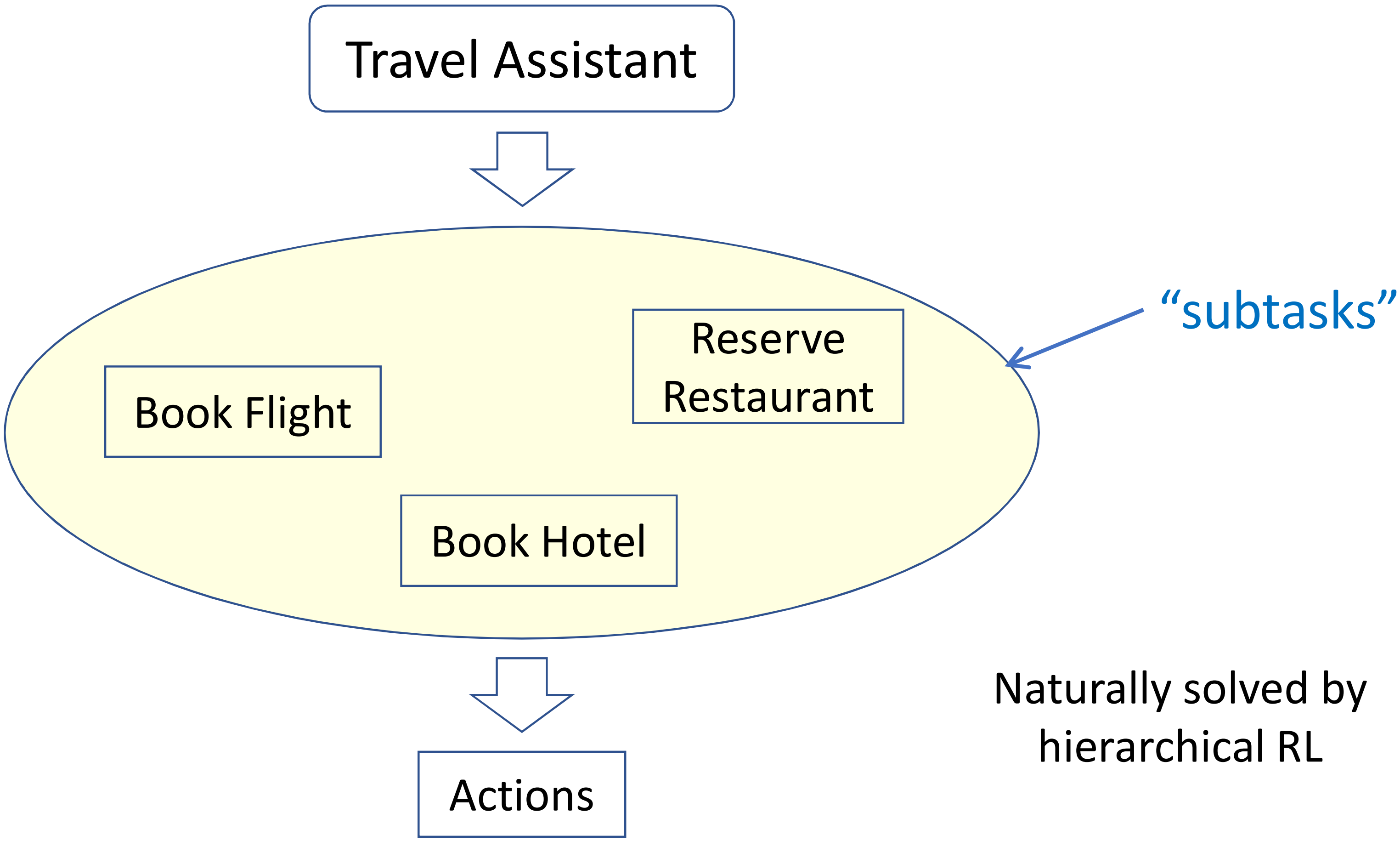
$$\hat{\theta} = \arg\min_{\theta_L} \mathrm{KL}\big(q(\mathbf{w}|\theta_L) \,||\, p(\mathbf{w}|Data)\big)$$

Still use "target network" θ_T
to synthesize regression target

- Parameter learning: solve for $\hat{\theta}$ with Bayes-by-backprop [Blundell et al. 2015]
- Params θ quantifies uncertainty in Q-values
- Action selection: use Thompson sampling for exploration

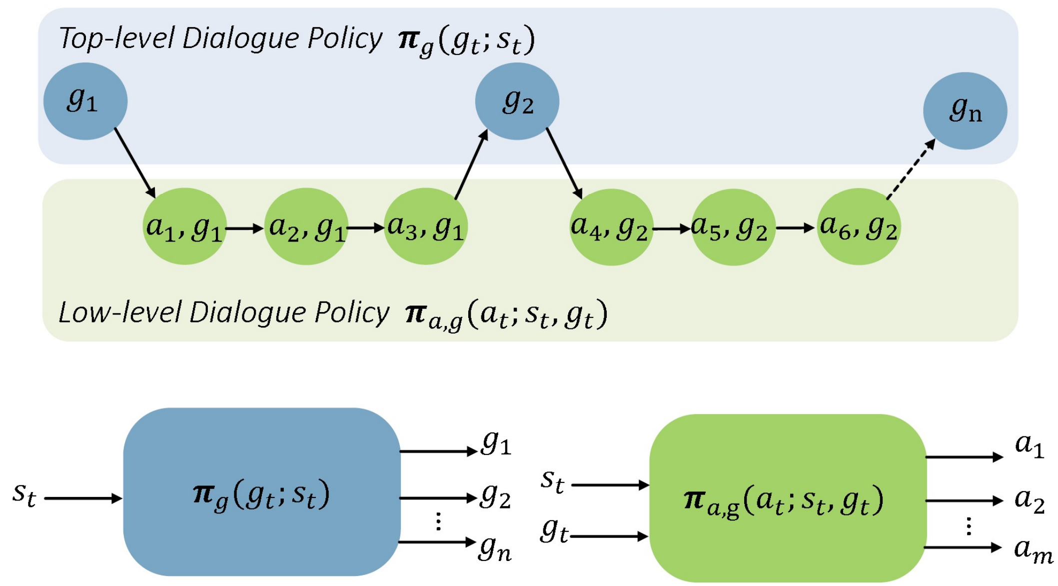
[Lipton+ 18]

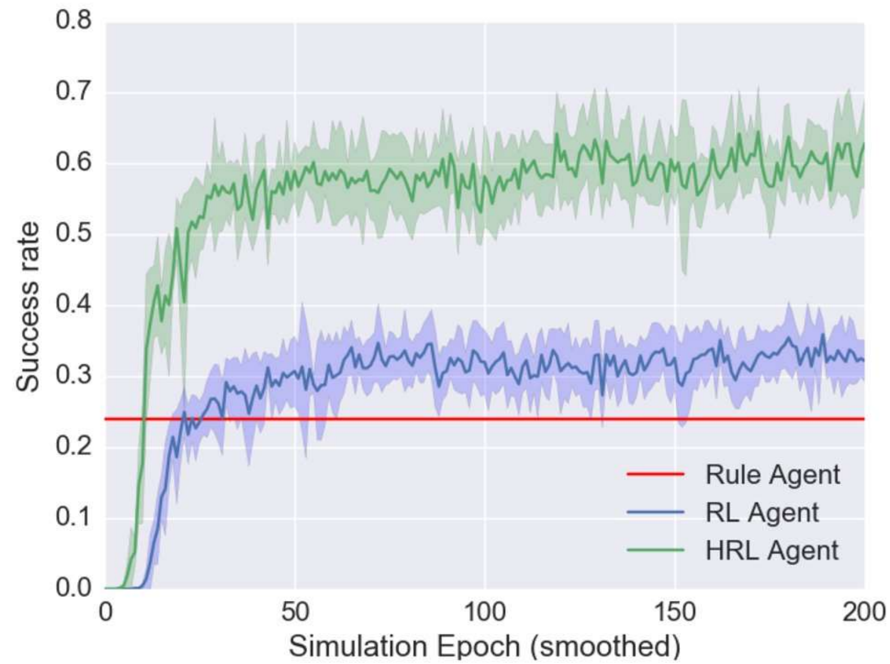
Composite-task Dialogues



Naturally solved by hierarchical RL

A Hierarchical Policy Learner

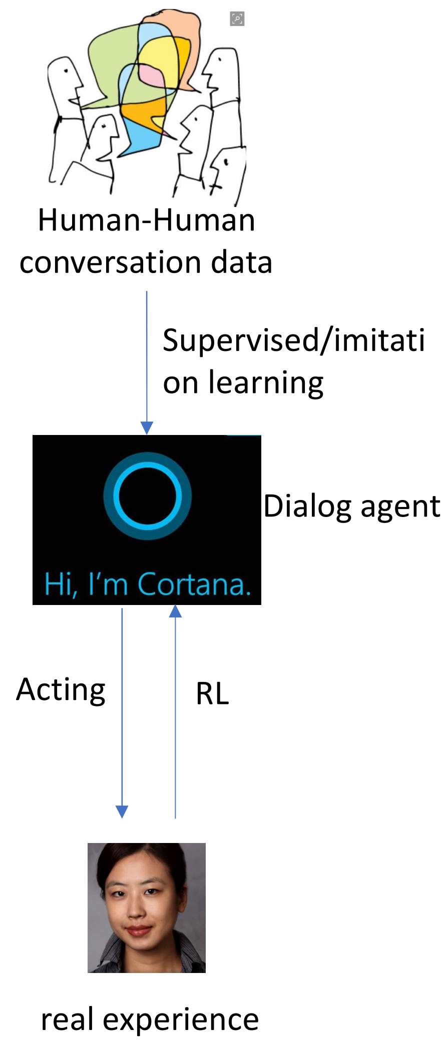




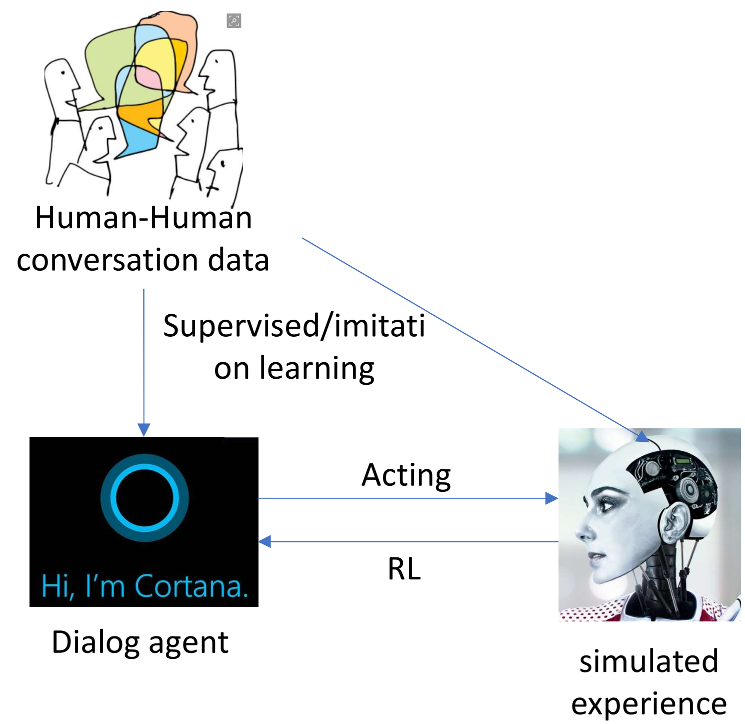
Similar to Hierarchical Abstract Machine (HAM) [Parr'98]

Superior results in both simulated and real users [Peng+ 17]

- Expensive: need large amounts of real experience except for very simple tasks
- Risky: bad experiences (during exploration) drive users away



- Inexpensive: generate large amounts of simulated experience for free
- Overfitting: discrepancy btw real users and simulators



Integrating Planning for Dialogue Policy Learning [Peng+ 18]

Dialogued agent trained using

- Limited real user experience
- Large amounts of simulated experience

Limited real experience is used to improve

- Dialog agent
- Planner (simulated user)

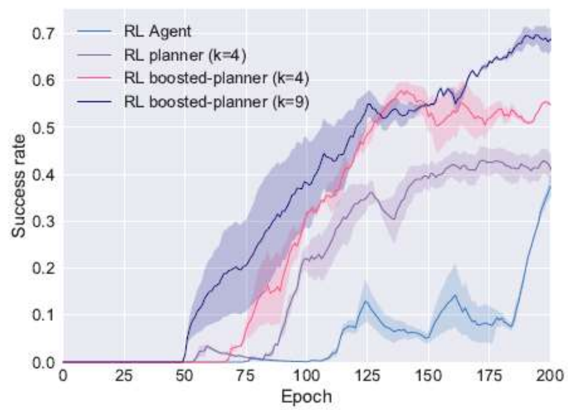
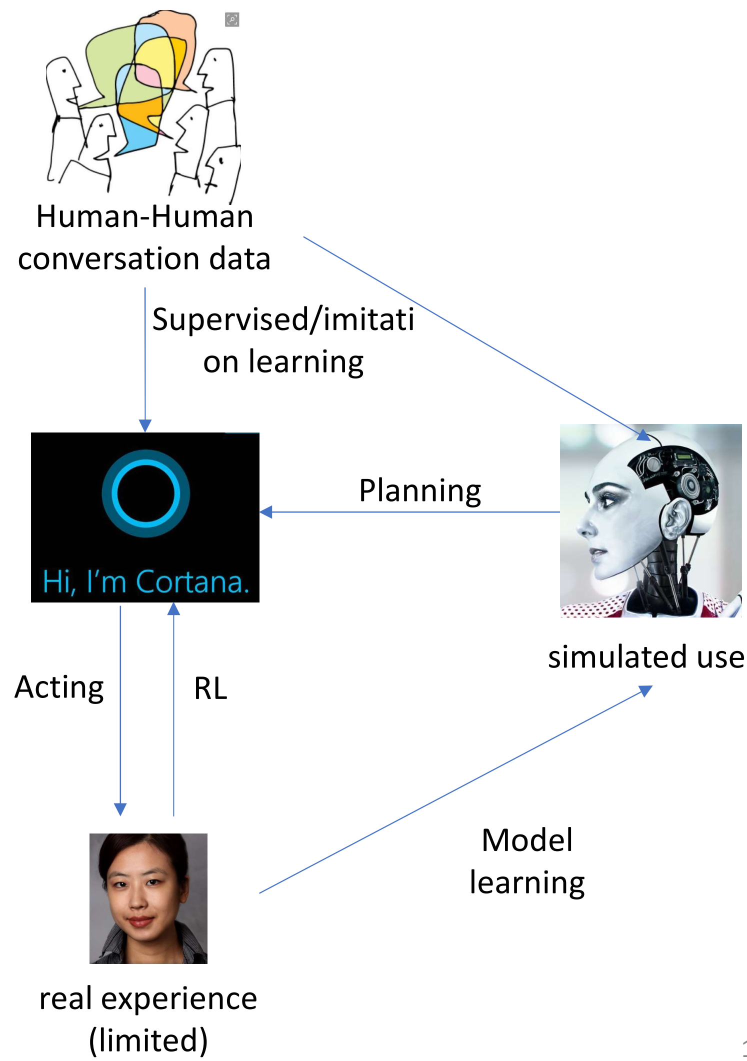


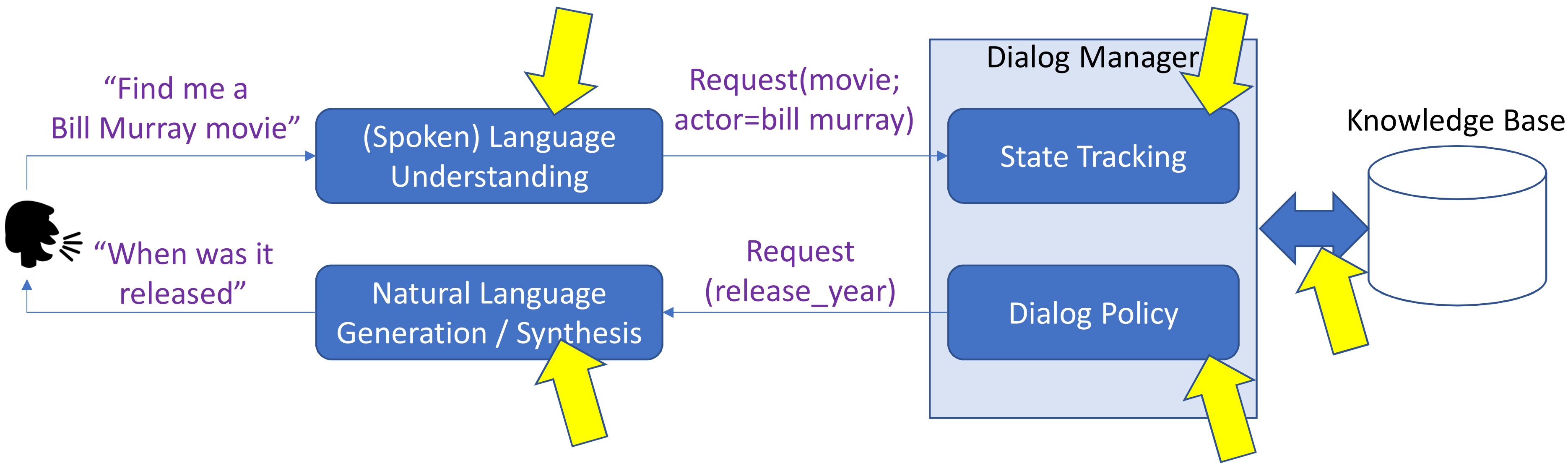
Figure 6: Human-in-the-loop Dialogue Policy Learning Curves for four different agents: x-axis is the number of training epochs.



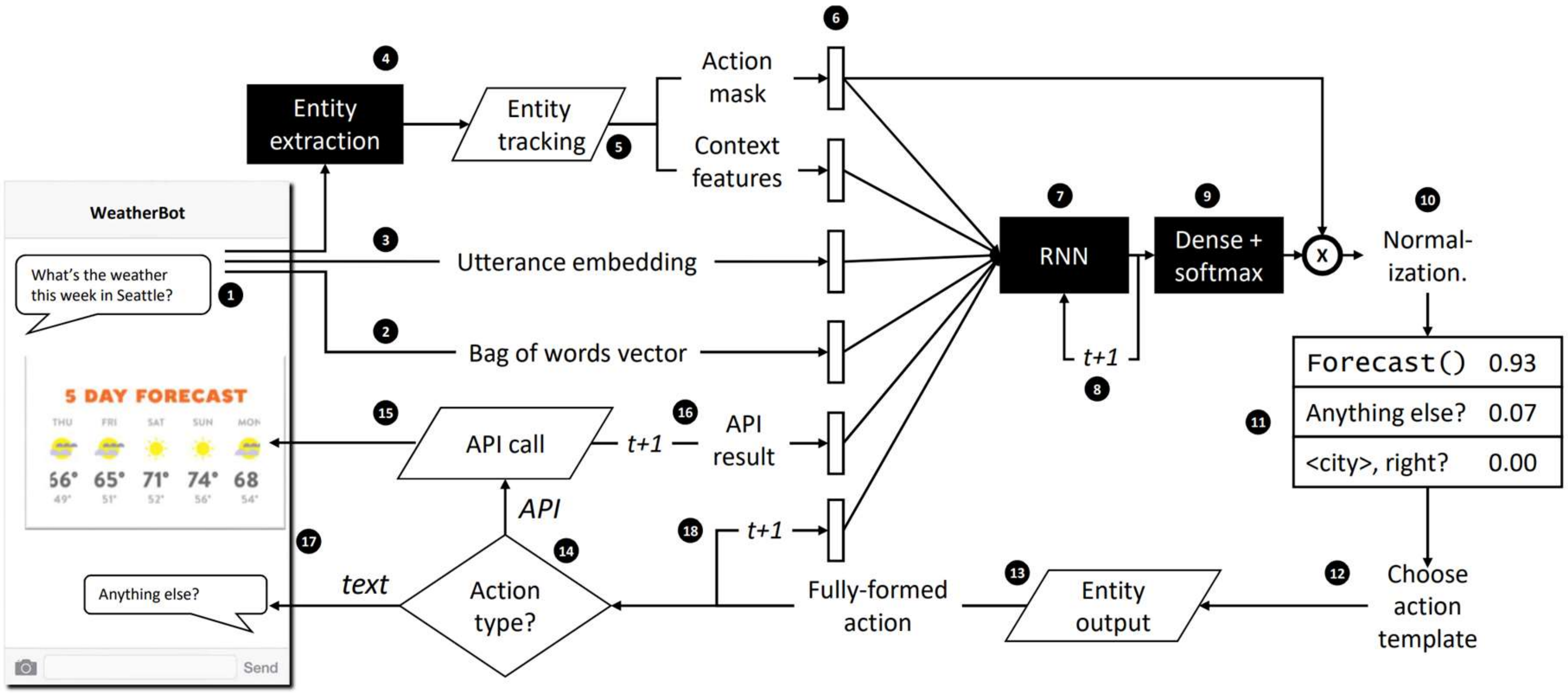
Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- **Part 3: Task-oriented dialogue**
  - Overview
  - Traditional approaches
  - Natural language understanding and dialogue state tracking
  - Deep RL for dialogue policy learning
  - **Neural approaches and E2E learning**
- Part 4: Fully data-driven conversation models and chatbots

Decision-theoretic View of Dialogue Management

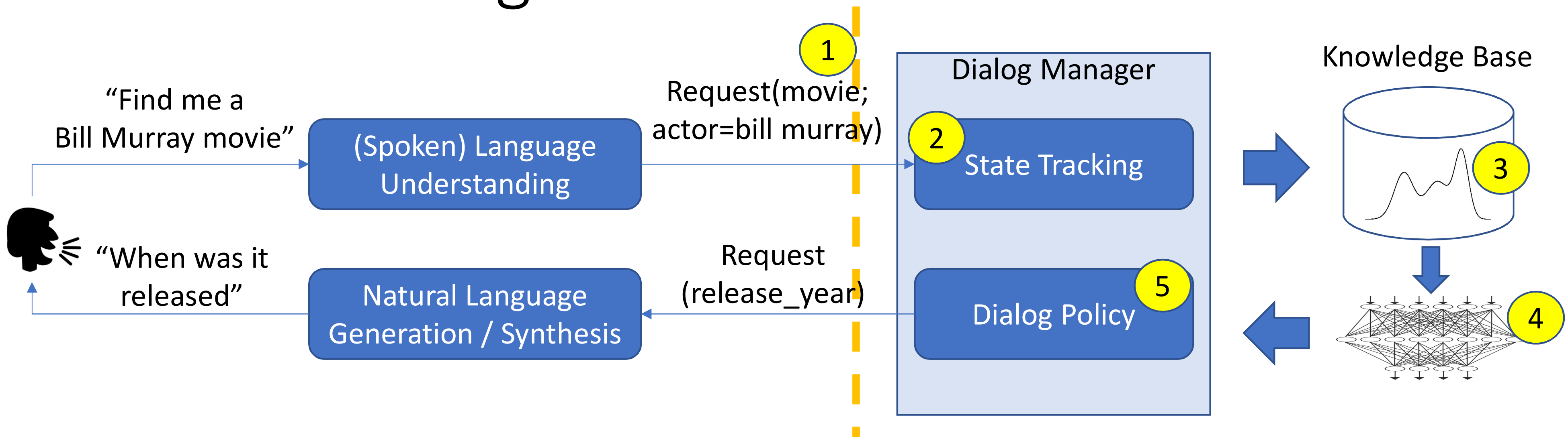


Hybrid Code Networks [Williams+ 17]



Allows domain knowledge to be encoded as API calls or action templates

Differentiating KB Accesses [Dhingra+ 17]



1. Use a single deep NN for {dialog manager *and* KB}
2. Recurrent network to track states of conversation
3. **Maintain (implicitly) a distribution over entities in KB**
4. A summary network to "summarize" distribution information
5. Multilayer perceptron policy network

Whole network can be end-to-end trained by BP/SGD!

Soft KB-lookup

Entity-Centric Knowledge Base

Movie	Actor	Release Year
Groundhog Day	Bill Murray	1993
Australia	Nicole Kidman	X
Mad Max: Fury Road	X	2015

- Posterior computation:

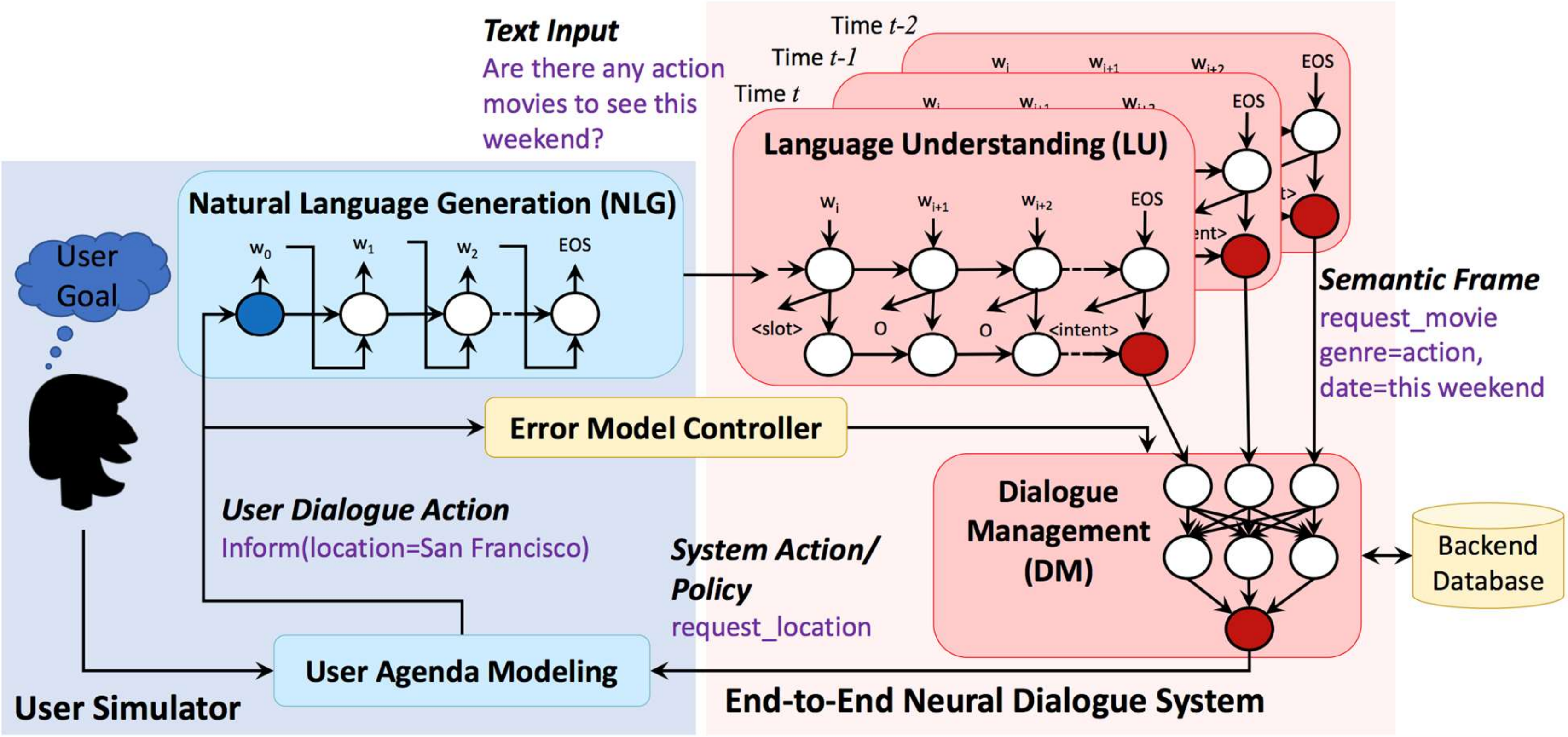
$$\mathrm{Pr}(\text{"GroundhogDay"}) \propto \mathrm{Pr}(\text{Actor} = \text{"Bill Murray"}) \cdot \mathrm{Pr}(\text{ReleaseYear} = \text{"1993"}) \cdots$$

Each $\mathrm{Pr}(\text{slot} = \text{value})$ is computed in terms of LU outputs

- Soft KB-lookup: sample a movie according to the posterior
  - Randomization results in differentiability (similar to policy gradient alg.)
  - As opposed to using SQL queries to look up results deterministically

Whole system can be trained using policy gradient & back-propagation

An E2E Neural Dialogue System [Li+ 17]



A more empirical study of end-to-end dialogue systems

Microsoft Dialogue Challenge at SLT-2018

- Task: build E2E task-oriented dialogue systems
- Data: labeled human conversations in 3 domains
- Experiment platform with built-in user simulators for training and evaluation
- Final evaluation in simulated setting and by human judges

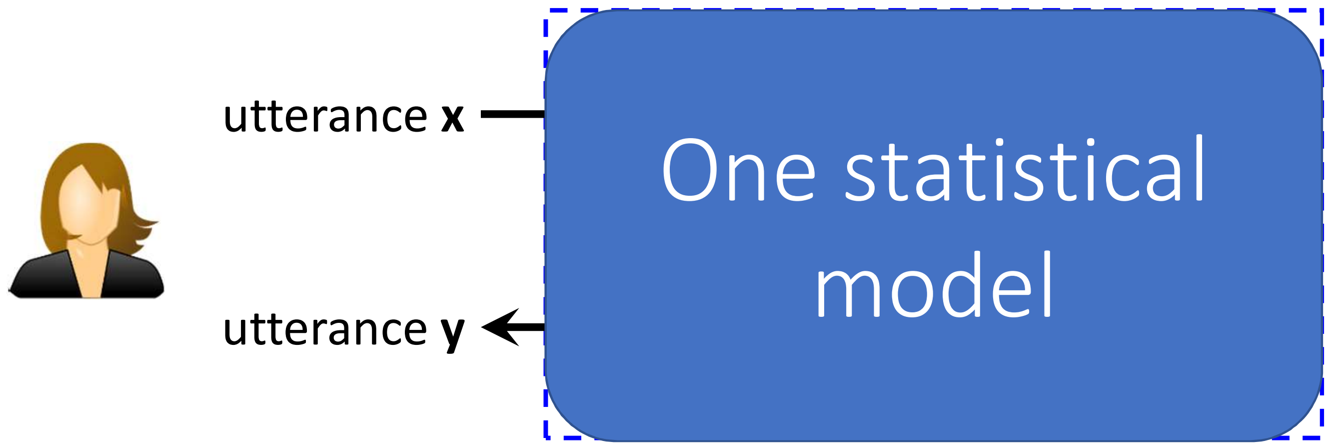
- More information:

  https://github.com/xiul-msr/e2e_dialog_challenge

Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
    - **E2E neural conversation models**
    - **Challenges and remedies**
    - **Grounded conversation models**
    - **Beyond supervised learning**
    - **Data and evaluation**
    - **Chatbots in public**
    - **Future work**

Motivation



Move towards **fully data-driven**, **end-to-end** dialogue systems.

Social Bots

- Fully end-to-end systems so far most successfully applied to **social bots** or **chatbots**:
  - Commercial systems: Amazon Alexa, XiaoIce, etc.

- Why social bots?
  - Maximize **user engagement** by generating **enjoyable** and **more human-like** conversations
  - Help **reduce user frustration**
  - **Influence dialogue research** in general (social bot papers often cited in task-completion dialogue papers)

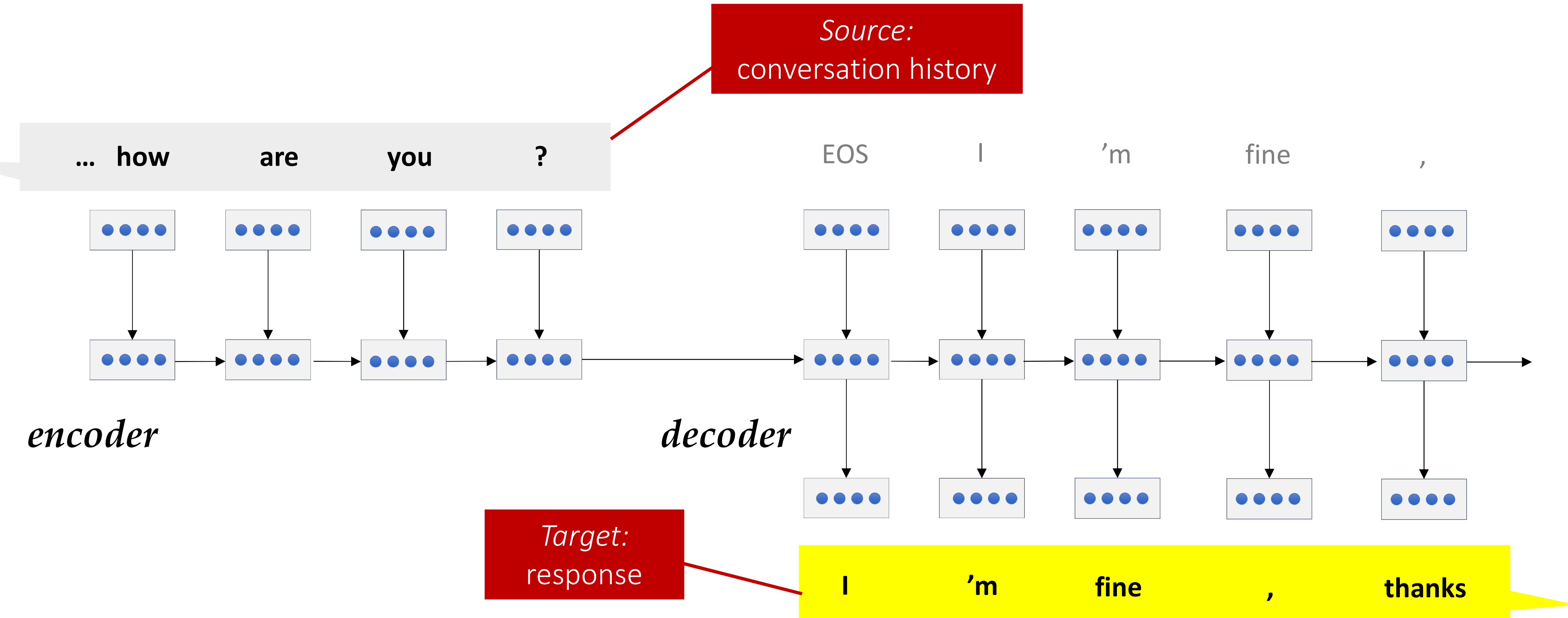
Historical overview

Earlier work in fully data-driven response generation:

- **2010:** Response retrieval system (IR) [Jafarpour+ 10]
- **2011**: Response generation using Statistical Machine Translation (phrase-based MT) [Ritter+ 11]
- **2015**: First neural response generation systems (RNN, seq2seq) [Sordoni+ 15; Vinyals & Le 15; Shang+ 15]

Neural Models for Response Generation

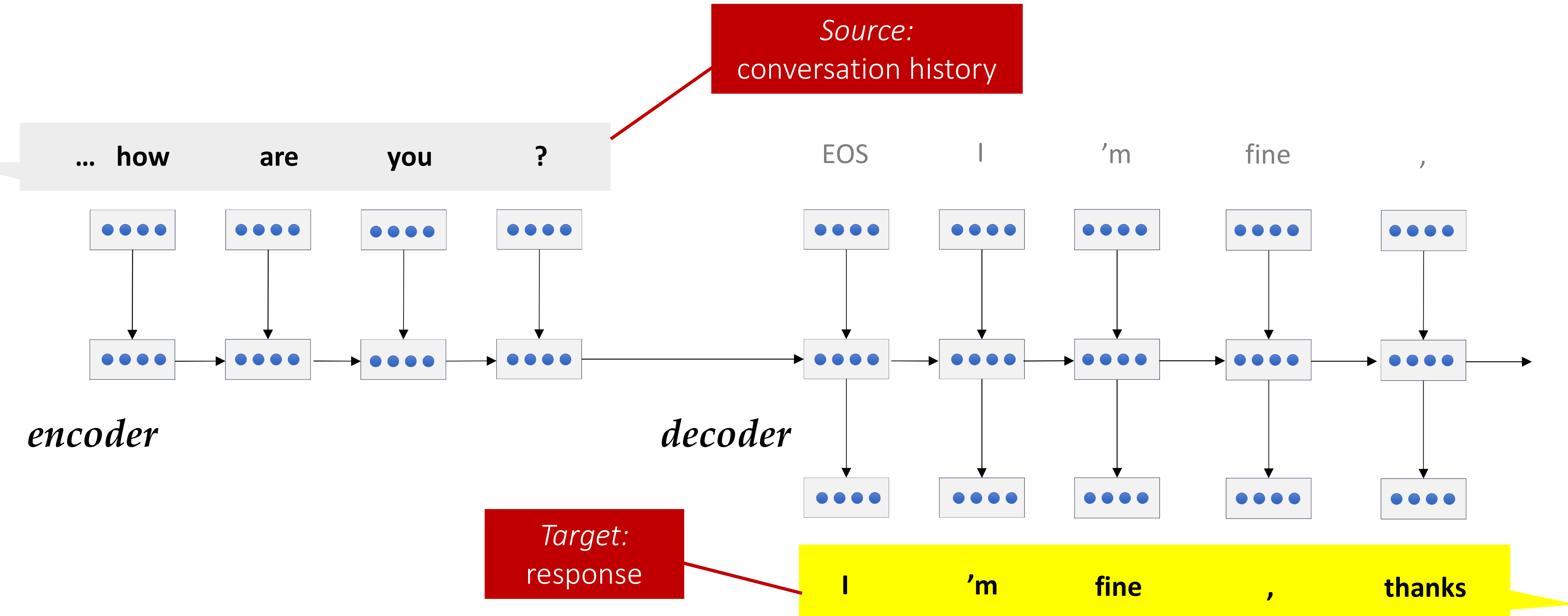
[Sordoni+ 15; Vinyals & Le 15; Shang+ 15]



Similar to sequence models in Neural Machine Translation (NMT), summarization, etc.
Uses either RNN, LSTM, GRU, etc.

Neural Models for Response Generation

[Sordoni+ 15; Vinyals & Le 15; Shang+ 15]



Neural Response Generation:
Difference with other tasks (e.g., machine translation)

- **Data:** some training sets (social media) are **HUGE**

  For example, Twitter (as of 2016):

  304M monthly active users

  500M tweets per day (6M conversations per day)

  → *often train on subset of the data; leaner and "faster" models preferred*

- **System input:** LONG conversation history

$T_1\ T_2\ T_3\ T_4 \ldots T_N$

Conversation (N turns)

$T_1 \rightarrow T_2$

$T_1\ T_2 \rightarrow T_3$

…

$\ldots T_{N-2}\ T_{N-1} \rightarrow T_N$

N-1 training instances

Either limit context size [Sordoni+ 05] or truncate backprop [Vinyals+ 05, pc]

[Sordoni+ 15; Li+ 16a]

Neural conversation engine

Neural model learns the general shape of conversation, and the system output is situationally appropriate and coherent…

H:	Hi
M:	**How are you?**
H:	Not bad
M:	**What's wrong?**
H:	Nothing really
M:	**Wanna talk about it? I'm here if you wanna talk x**
H:	Well, I'm just tired
M:	**Me too, I'm here if you wanna talk**
H:	Thanks, what are you doing?
M:	**You're welcome and I'm watching TV and you?**
H:	On Netflix?
M:	**Yeah, I'm watching it on Netflix.**
H:	What are you watching?
M:	**A TV show.**
H:	Nice, did you have dinner already?
M:	**No, not yet. what about you?**

A more goal-oriented example. No grounding into a real calendar, but the “shape” of the conversation is fluent and plausible…

H:	Would Thursday afternoon work sometime?
M:	**Yeah, sure. just let me know when you‘re free.**
H:	After lunch is probably the best time
M:	**Okay, sounds good. Just let me know when you’re free.**
H:	Would 2pm work for you?
M:	**Works for me.**
H:	Well let’s say 2pm then I’ll see you there
M:	**Sounds good.**

Neural E2E system trained on 35M Twitter conversations.

Fully Data-driven Response Generation:

Challenges and remedies

Challenge: The blandness problem

How was your weekend?

I don't know.

What did you do?

I don't understand what you are talking about.

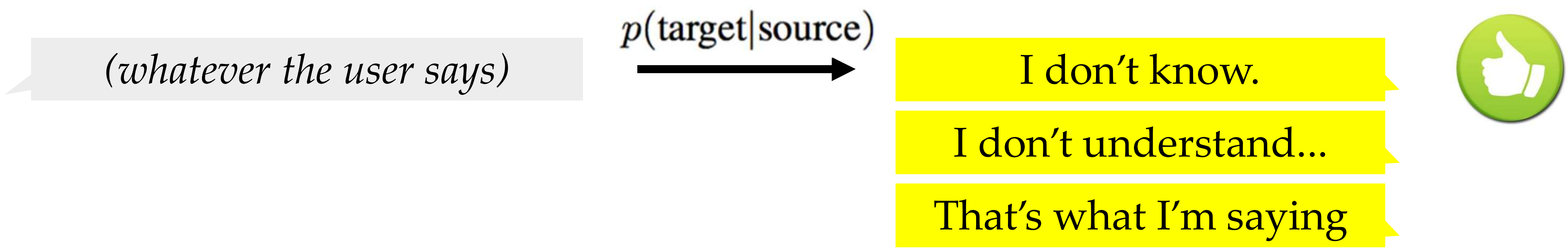
This is getting boring…

Yes that's what I'm saying.

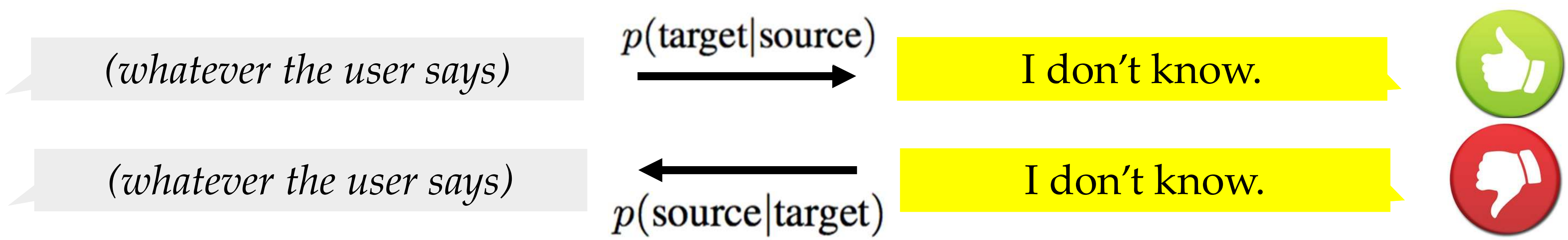
[[Li+ 16a]]

Blandness problem: cause and remedies

Common MLE objective (maximum likelihood)



Mutual information objective:



[Li+ 16a]

Mutual Information for Neural Network Generation

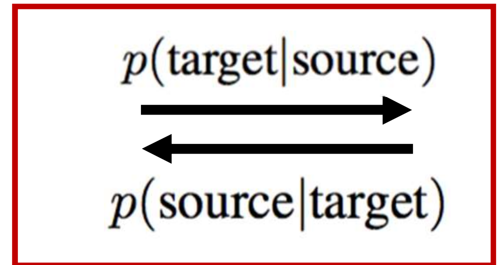
Mutual information objective:

$$\hat{T} = \arg\max_{T} \left\{ \log \frac{p(S,T)}{p(S)p(T)} \right\}$$

$$\hat{T} = \arg\max_{T} \left\{ \log p(T|S) - \lambda \log p(T) \right\}$$

standard likelihood

anti-LM



Bayes' rule

Bayes' theorem

$$\hat{T} = \arg\max_{T} \left\{ (1-\lambda) \log p(T|S) + \lambda \log p(S|T) \right\}$$

Sample outputs (MMI)

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

Of course you can! They're delicious!

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

I think he'd be a good signing.

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

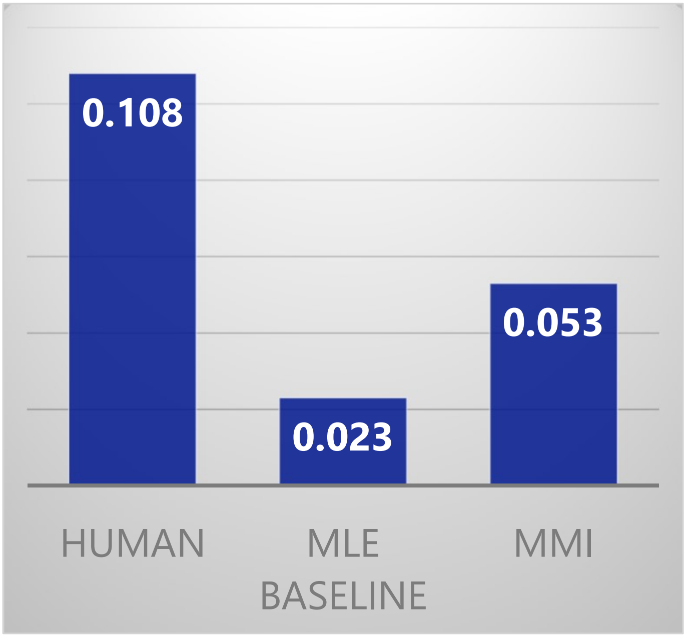
Depends on how much you drink!

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

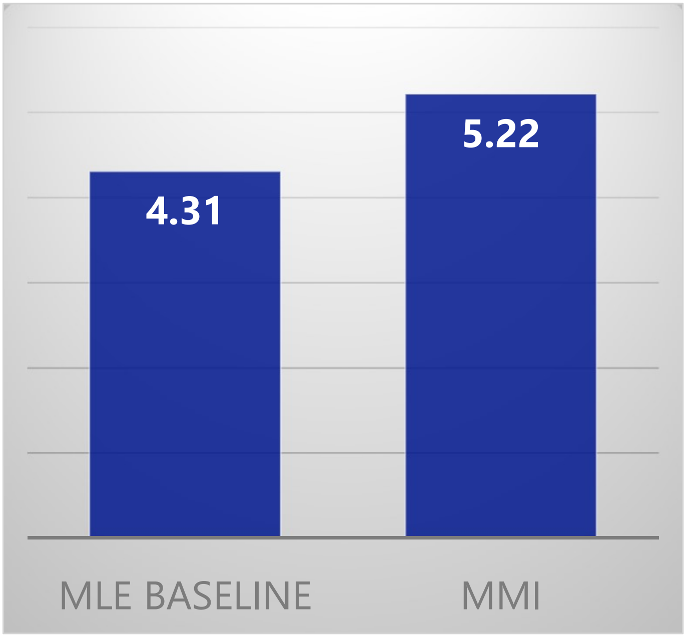
I've never seen him live.

[Li+ 16a]

MLE vs MMI: results



Lexical diversity
(# of distinct tokens / # of words)



BLEU

MMI: best system in Dialogue Systems Technology Challenge 2017 (**DSTC**, E2E track)

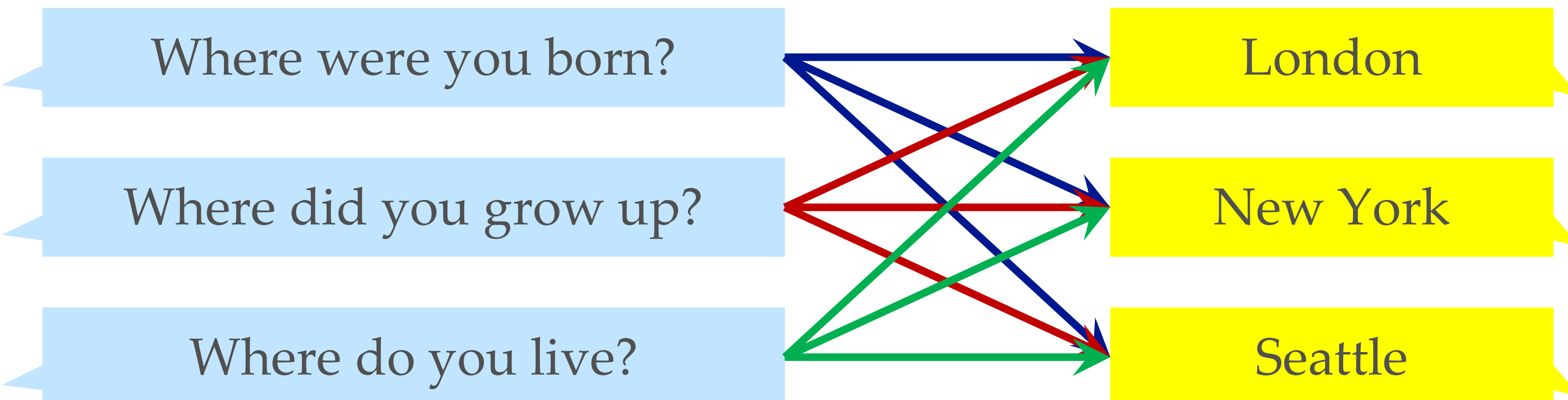
Challenge: The consistency problem

- E2E systems often exhibit **poor response consistency**:

message	Where is your hometown?
response	I live in Indonesia.
message	Where are you from?
response	England, you?
message	In which city do you live now?
response	I live in Jakarta.
message	In which country do you live now?
response	I live in Indonesia.
message	What is your major?
response	Business, you?
message	What did you study in college?
response	Psychology, you?
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 16.

The consistency problem: why?

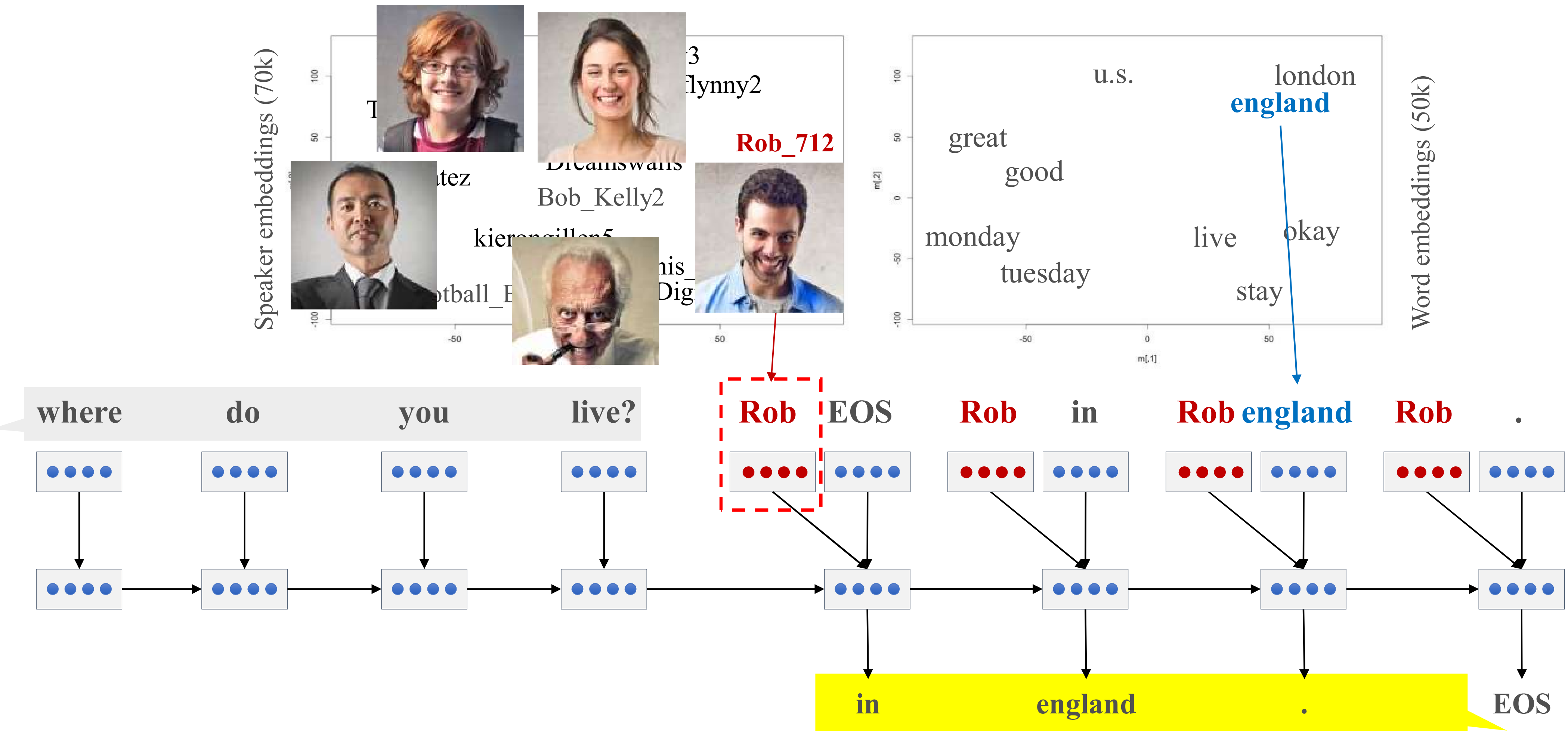
Conversational data:



NOT
1-to-1

P(response | query, **SPEAKER_ID**)

Personalized Response Generation [Li+ 2016b]



Persona model results

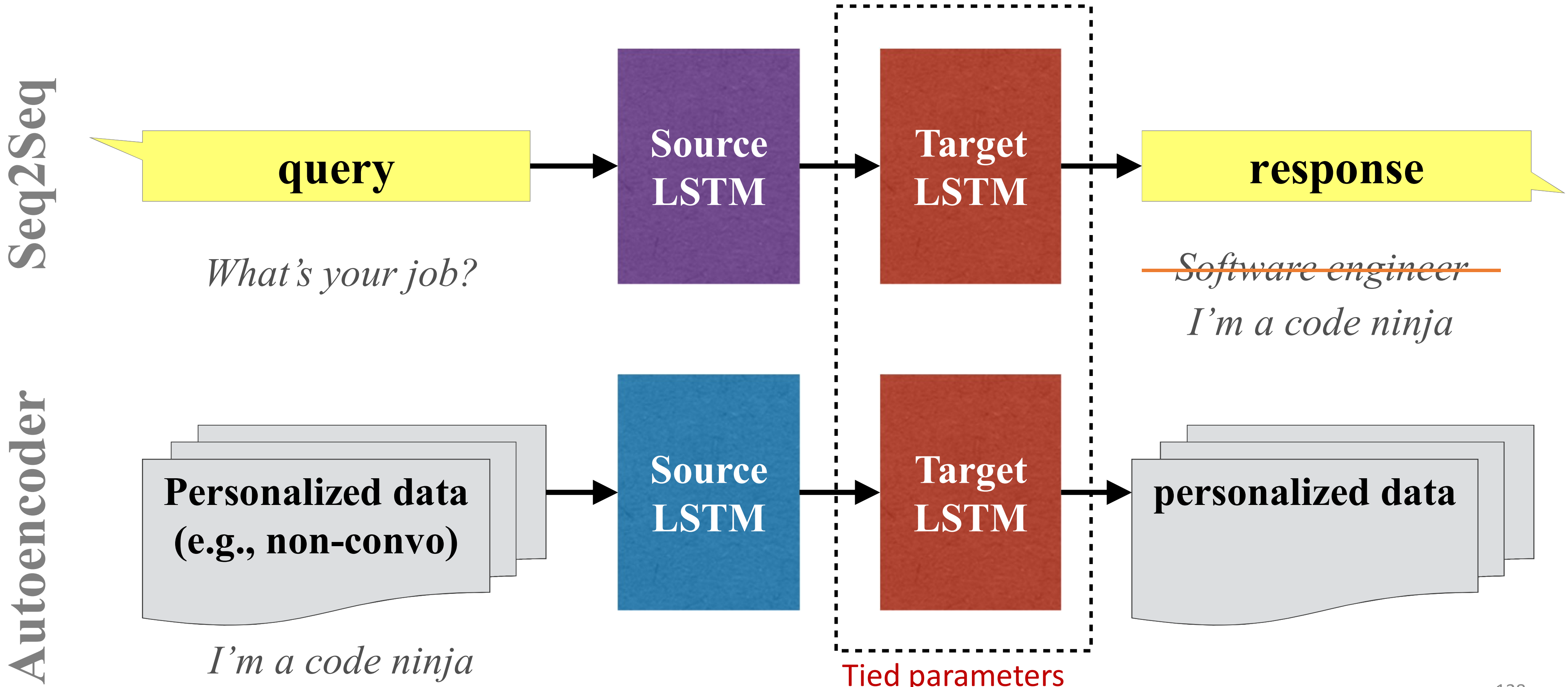
Baseline model:

message	Where is your hometown?
response	I live in Indonesia.
message	Where are you from?
response	England, you?
message	In which city do you live now?
response	I live in Jakarta.
message	In which country do you live now?
response	I live in Indonesia.
message	What is your major?
response	Business, you?
message	What did you study in college?
response	Psychology, you?
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 16.

Persona model using speaker embedding: [Li+ 16b]

message	Where is your hometown?
response	I'm from England.
message	Where are you from?
response	I'm from England.
message	In which city do you live now?
response	I live in London.
message	In which country do you live now?
response	I live in England.
message	What is your major?
response	Business. You?
message	What did you study in college?
response	I did business studies.
message	How old are you?
response	I'm 18.
message	What is your age?
response	I'm 18.

Personal modeling as multi-task learning [Luan+ 17]



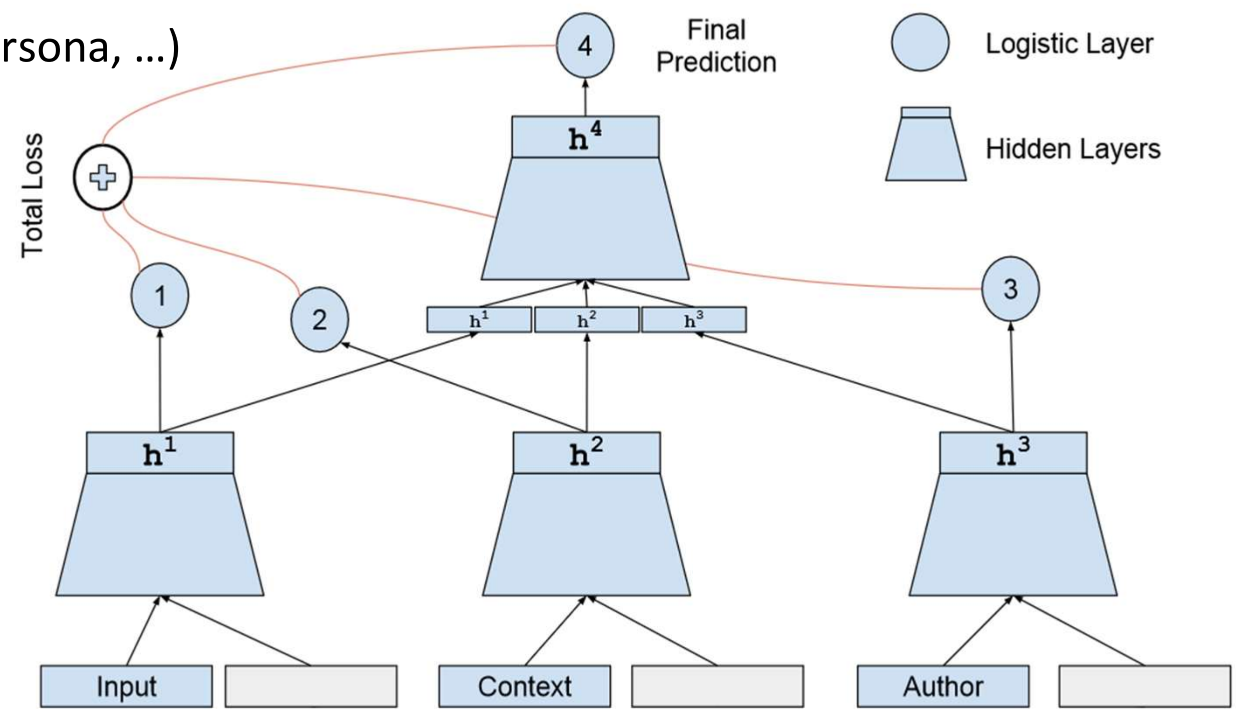
Improving personalization with multiple losses [Al-Rfou+ 16]

- **Single-loss:**
  P(response | context, query, persona, …)

  Problem with single-loss: context or query often **“explain away”** persona

- **Multiple loss adds:**
  P(response | persona)
  P(response | query)
  etc.

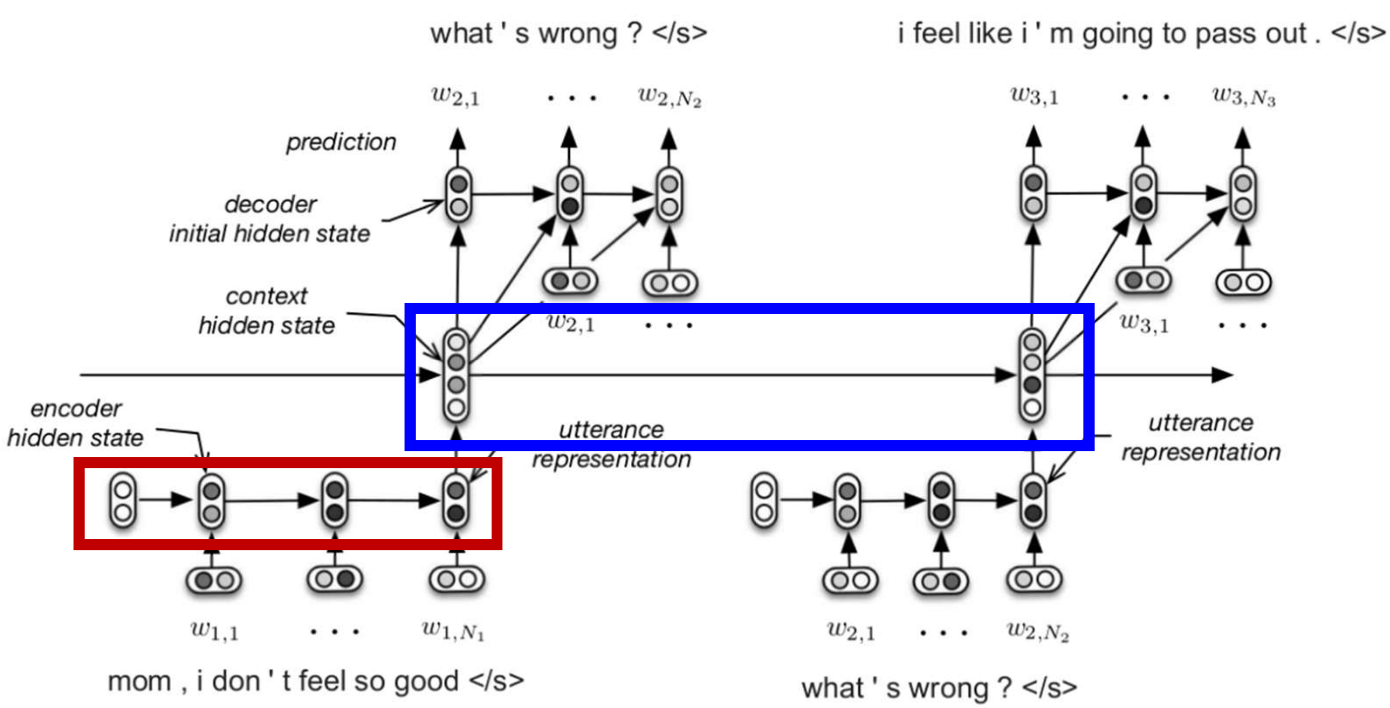
  Optimized so that persona can “predict” response all by itself → more robust speaker embeddings



Challenge: Long conversational context

It can be challenging for LSTM/GRU to encode very long context (i.e. more than 200 words: [Khandelwal+ 18])

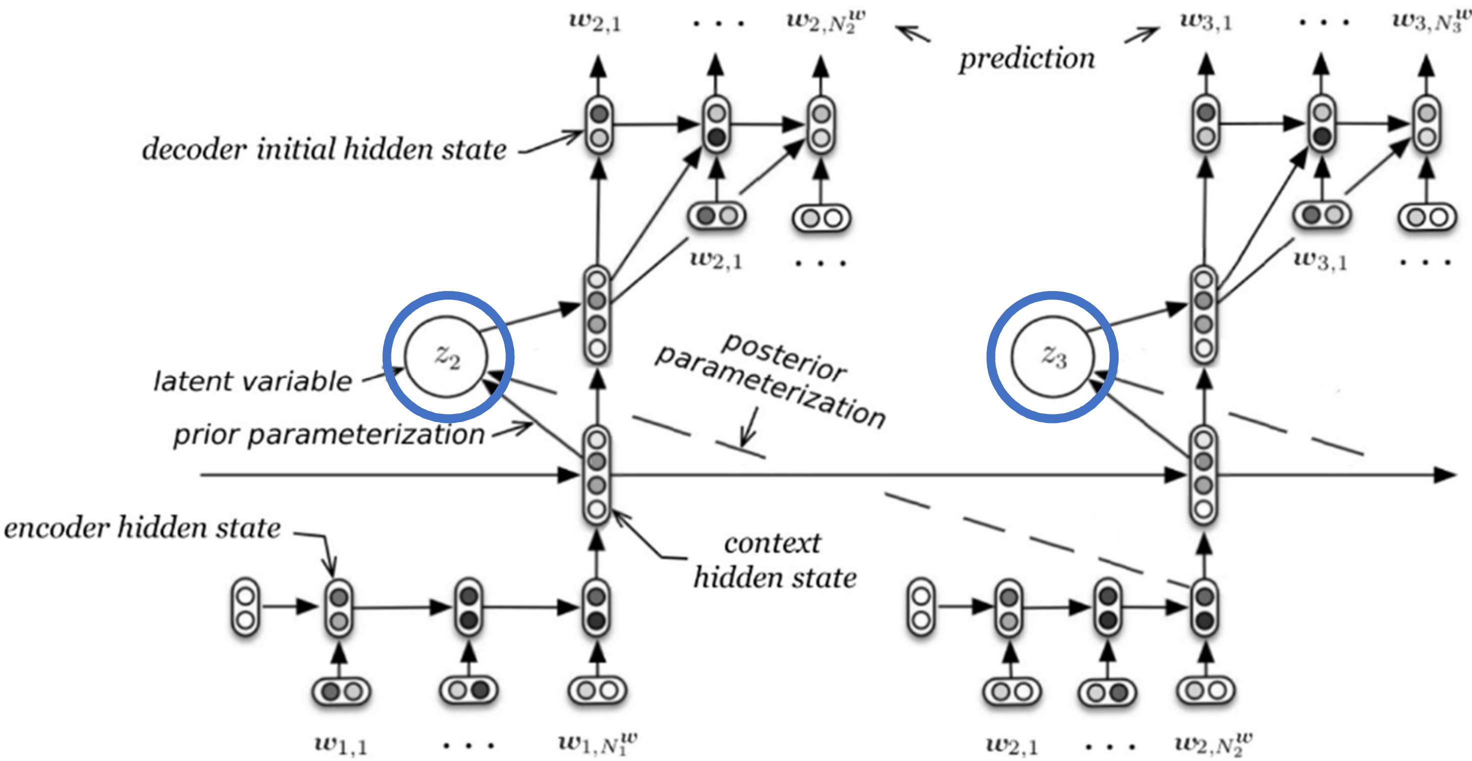
- **Hierarchical Encoder-Decoder (HRED)** [Serban+ 16]



Encodes:
utterance (word by word) +
conversation (turn by turn)

Challenge: Long conversational context

- **Hierarchical Latent Variable Encoder-Decoder (VHRED)** [Serban+ 17]
  - Adds a latent variable to the decoder
  - Trained by maximizing variational lower-bound on the log-likelihood



Related to persona model [Li+ 2016b]:

Deals with 1-N problem, but unsupervisedly.

Hierarchical Encoders and Decoders: Evaluation

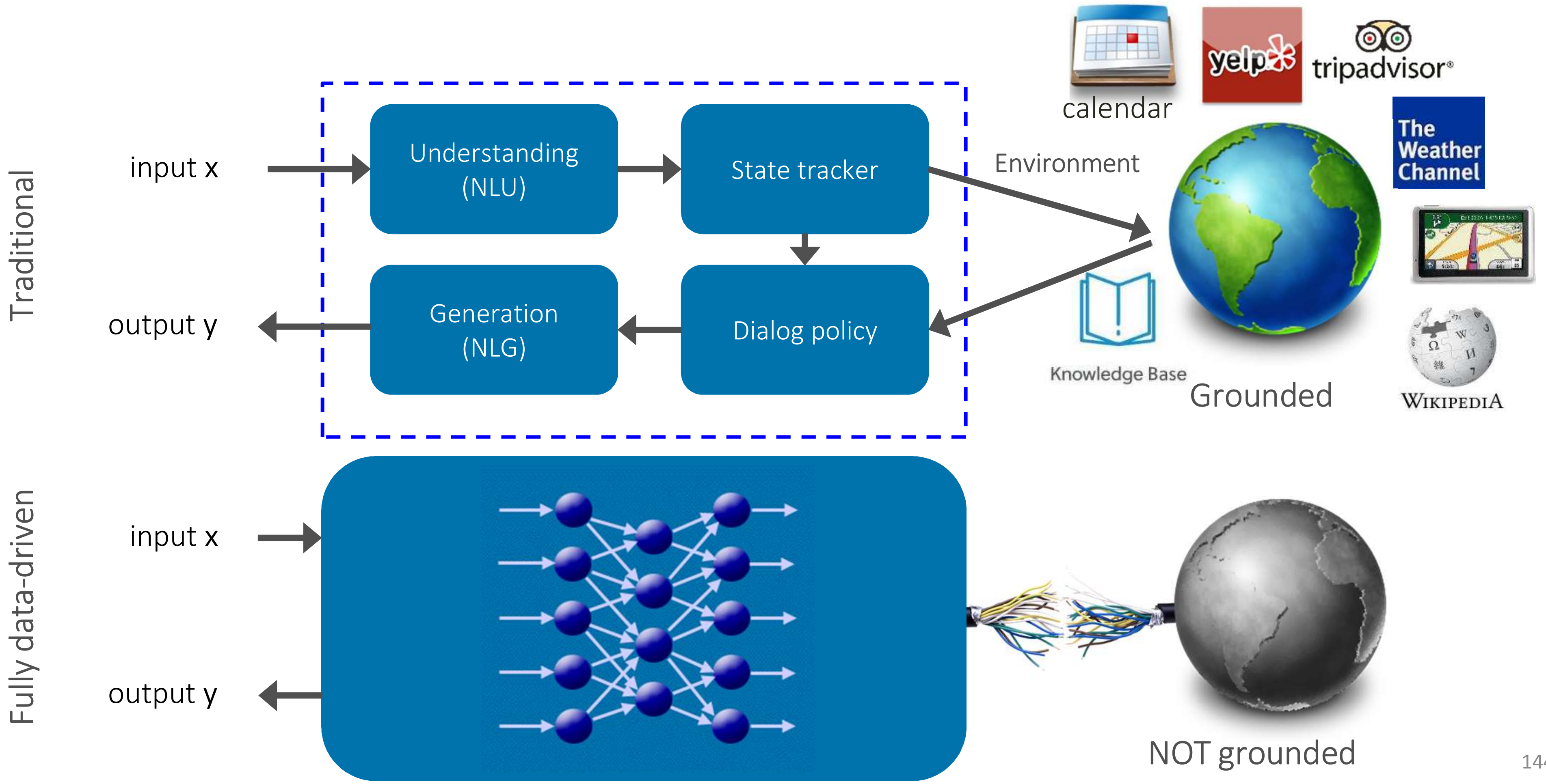
[Serban+ 17]

Opponent	Wins	Losses	Ties
Short Contexts			
VHRED vs LSTM	32.3 $\pm$2.4	**42.5 $\pm$2.6***	25.2 $\pm$2.3
VHRED vs HRED	**42.0 $\pm$2.8***	31.9 $\pm$2.6	26.2 $\pm$2.5
VHRED vs TF-IDF	**51.6 $\pm$3.3***	17.9 $\pm$2.5	30.4 $\pm$3.0
Long Contexts			
VHRED vs LSTM	**41.9 $\pm$2.2***	36.8 $\pm$2.2	21.3 $\pm$1.9
VHRED vs HRED	**41.5 $\pm$2.8***	29.4 $\pm$2.6	29.1 $\pm$2.6
VHRED vs TF-IDF	**47.9 $\pm$3.4***	11.7 $\pm$2.2	40.3 $\pm$3.4

Outline

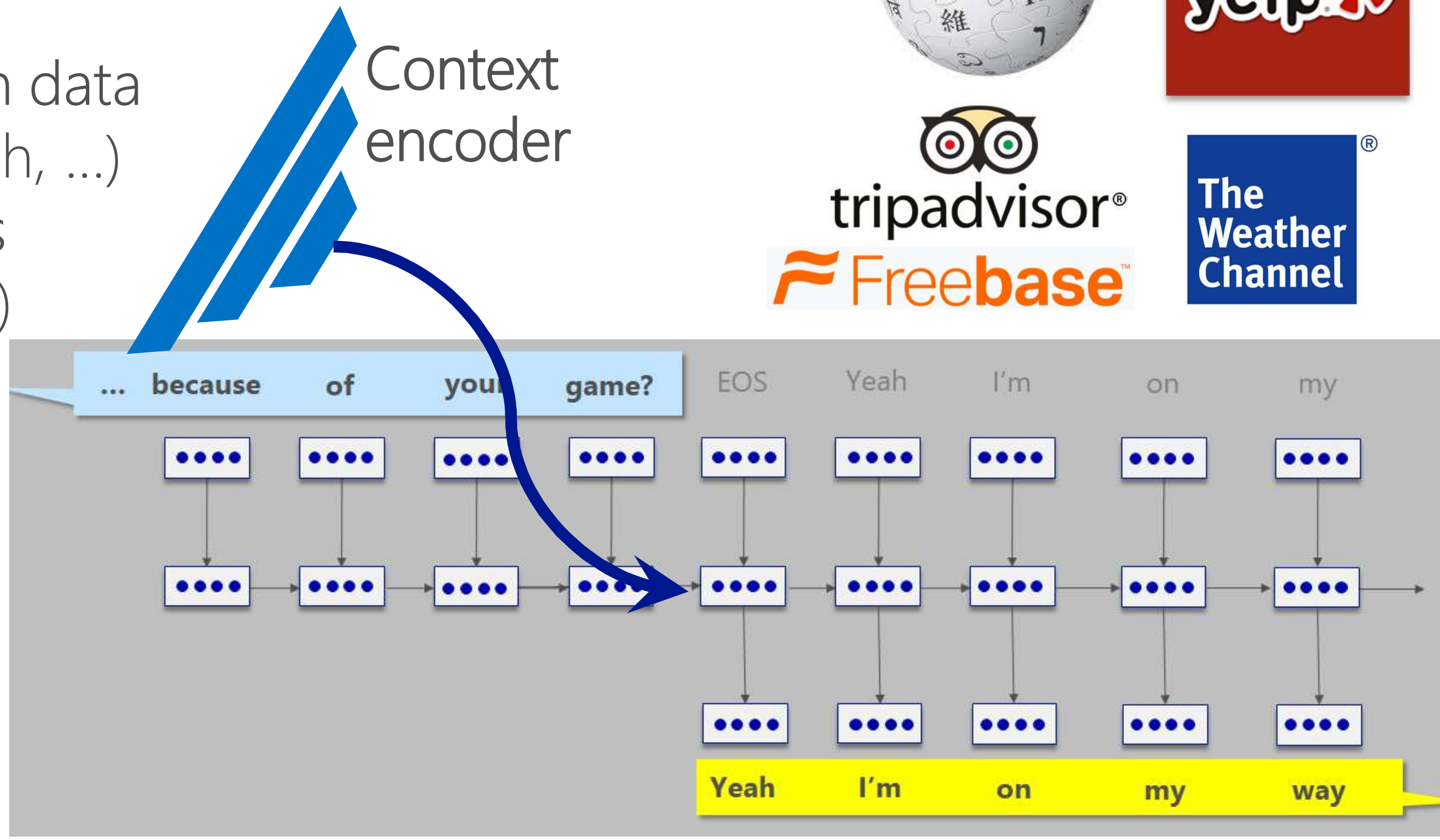
- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
    - **E2E neural conversation models**
    - **Challenges and remedies**
    - **Grounded conversation models**
    - **Beyond supervised learning**
    - **Data and evaluation**
    - **Chatbots in public**
    - **Future work**

Towards Grounded E2E Conversation Models



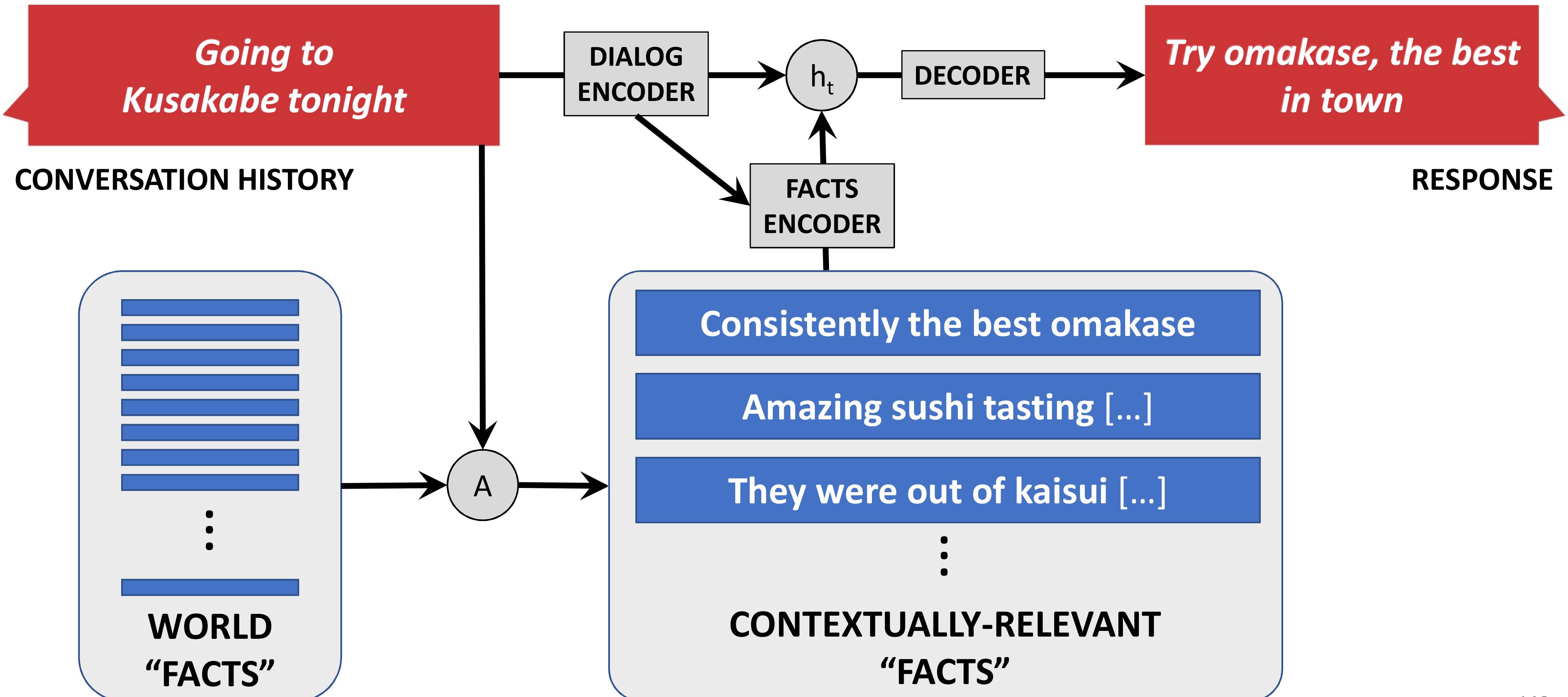
E2E Conversation Models in the real world

Personalization data
(ID, social graph, …)
Device sensors
(GPS, vision, …)
External
"knowledge"



A Knowledge-Grounded Neural Conversation Model

[Ghazvininejad+ 17]



Learns implicit slot filling

You know any good **A** restaurant in **B**?

Try **C**, one of the best **D** in the city.

You know any good Japanese restaurant in Seattle?

Try Kisaku, one of the best sushi restaurants in the city.

Sample knowledge-grounded responses

FOURSQUARE

I would give ALMOST anything for some [*Mexican restaurant*] right now.

Me too. Jalapeno sauce is really good.

Obsessed with [*jewelry company*] :-*

oh my gosh obsessed with their bracelets and the meaning behind them!

Visiting the celebs at Los Angeles airport - [...] w/ 70 others

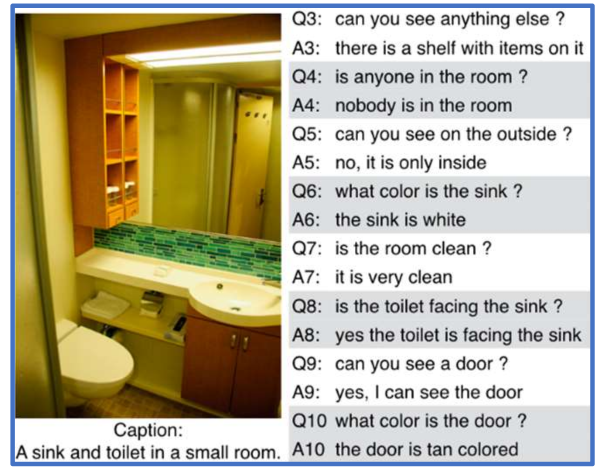
Nice airport terminal. Have a safe flight.

Experimental results (23M conversations): outperforms competitive neural baseline (human + automatic eval)

Grounded E2E Dialogue Systems

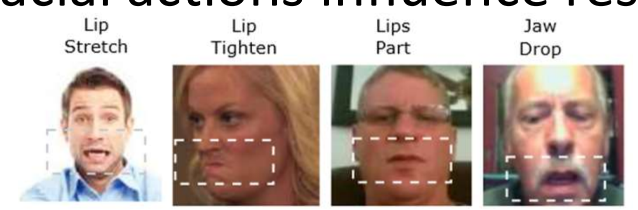
- **Grounding: images**

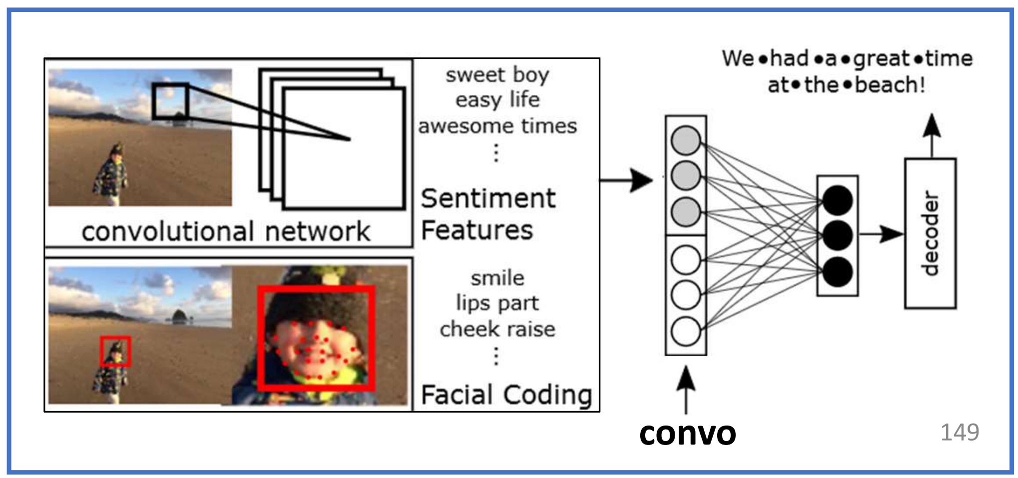
  Conversations around images e.g.,
  Q-As [Das+ 16] or chat [Mostafazadeh+ 17]



- **Grounding: affect** [Huber+ 18]

  facial actions influence response





Beyond supervised learning

- **Limitations of SL for E2E dialogue:**
  - Train on human-human data, test with human-machine (Twitter-ese often not what we want at test time.)
  - Optimizes for immediate reward $p(T_N \mid \ldots T_{N-1})$, not long-term reward
  - No user feedback loop

- **Emergence of reinforcement learning (RL) for E2E dialogue**
  - Tries to promote long-term dialogue success

Deep Reinforcement Learning for E2E Dialogue

[Li+ 16c]

- REINFORCE algorithm [Williams+ 92]

$$J(\theta) = \mathbb{E}[R(s_1, s_2, ..., s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) \boxed{R(s_1, s_2, ..., s_N)}$$

reward function

$$\nabla J(\theta) = \nabla \log \prod_i \boxed{p(s_i|s_{i-1})} R(s_1, s_2, ..., s_N)$$

what we want to learn

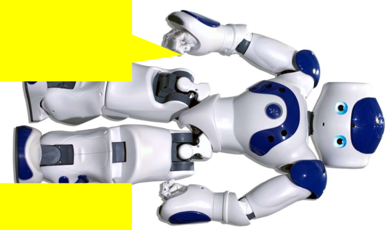
- Reward functions:
  1. Ease of answering: $-Pr(\text{Dull Response}|s_i)$
  2. Information flow: $-\log \textbf{Sigmoid} \cos(s_1, s_2)$
  3. Meaningfulness: $\log p_{\text{seq2seq}}(s_1|s_2) + \log p_{\text{seq2seq}}(s_2|s_1)$

Simulation (without RL)

See you later!

See you later!

See you later!



See you later!

See you later!

See you later!

152

Simulation (with RL)

How old are you ?

i 'm 4, why are you asking ?

I thought you were 12 .

What made you think so ?

I don’t know what you are talking about .

You don’t know what you are saying.

I don’t know what you are talking about .

Deep RL: Evaluation

- MTurk evaluation (500 responses)

Setting	RL-win	RL-lose	Tie
single-turn general quality	0.40	0.36	0.24
single-turn ease to answer	0.52	0.23	0.25
multi-turn general quality	0.72	0.12	0.16

Outline

- Part 1: Introduction
- Part 2: Question answering and machine reading comprehension
- Part 3: Task-oriented dialogue
- **Part 4: Fully data-driven conversation models and chatbots**
  - **E2E neural conversation models**
  - **Challenges and remedies**
  - **Grounded conversation models**
  - **Beyond supervised learning**
  - **Data and evaluation**
  - **Chatbots in public**
  - **Future work**

Conversational datasets
(for social bots, E2E dialogue research)

- Survey on dialogue datasets [Serban+ 15]

Name	Type / Topics	Size
Reddit	**Unrestricted**	**N/A (growing)**
Twitter	**Unrestricted**	**N/A (growing)**
OpenSubtitles	**Movie subtitles**	**1B words**
Ubuntu Dialogue Corpus	Chat on Ubuntu OS	100M words
Ubuntu Chat Corpus	Chat on Ubuntu OS	2B words

Evaluating E2E Dialogue Systems

- **Human evaluation (crowdsourcing):**

Context: … *Because of your game?*
Input: *Yeah, I'm on my way now*
Response: **Ok good luck!**

	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree
Is **this** a good[1] response?	○	○	○	●	○

[1]: replaced as appropriate (relevant, interesting,…)

- **Automatic evaluation:**
  Less expensive, but is it reliable?

Machine-Translation-Based Metrics

- **BLEU** [Papineni+ 02]: ngram overlap metric

Reference:	John	resigned	yesterday	.
System:	Yesterday ,	John	quit .	

$$\text{BLEU} = \text{BP} \cdot \exp\left(\sum_n \log p_n\right)$$

$$p_n = \frac{\sum_i \sum_{g \in n\text{-grams}(h_i)} \max_j \{\#_g(h_i, r_{i,j})\}}{\sum_i \sum_{g \in n\text{-grams}(h_i)} \#_g(h_i)}$$

- **NIST** [Doddington+ 02]
  - Seldom used in dialogue, but copes with **blandness issue**
  - Considers info gain of each ngram: score(*interesting calculation*) >> score(*of the*)
- **METEOR**
  - Accounts for synonyms, paraphrases, etc.

The challenge with MT-based metrics

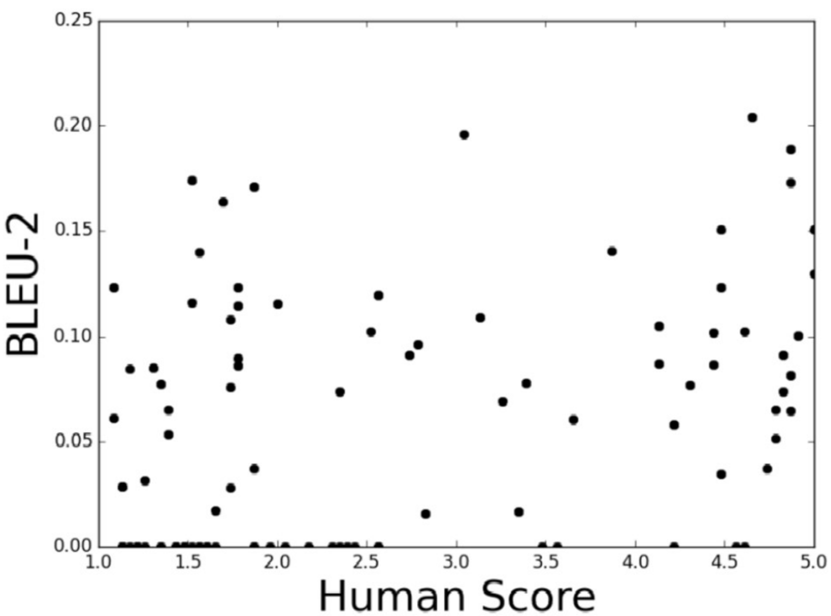
Input:	*How are you?*
Response (gold):	**I ’m good , thanks .**

Response A:	**Good thanks** !	***Semantically** equivalent (as in Machine Translation)*
Response B:	Doing pretty **good thanks**	
Response C:	Doing well thank you !	
Response D:	Fantastic **.** How are you ?	***Pragmatically** appropriate*
Response E:	**I 'm** getting sick again **.**	
Response F:	Bored **.** you ?	
Response G:	Sleepy **.**	
Response H:	Terrible tbh	

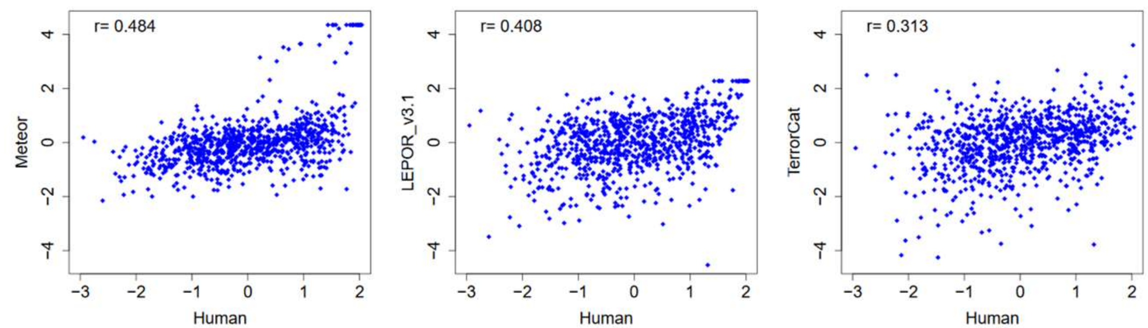
Many false negative!

Sentence-level correlation of MT metrics

- Poor correlation with human judgments:



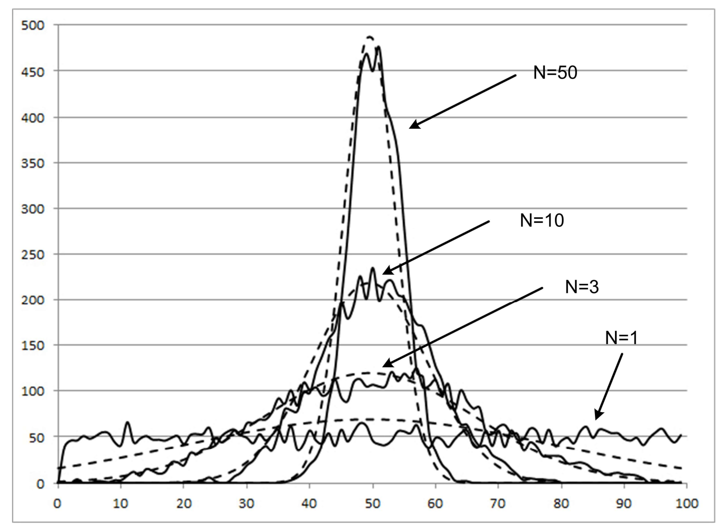
Dialogue task
"How NOT to evaluate dialogue systems"
[Liu+ 16]



But same problem even
for **Translation task**
[Graham +15]

The importance of sample size

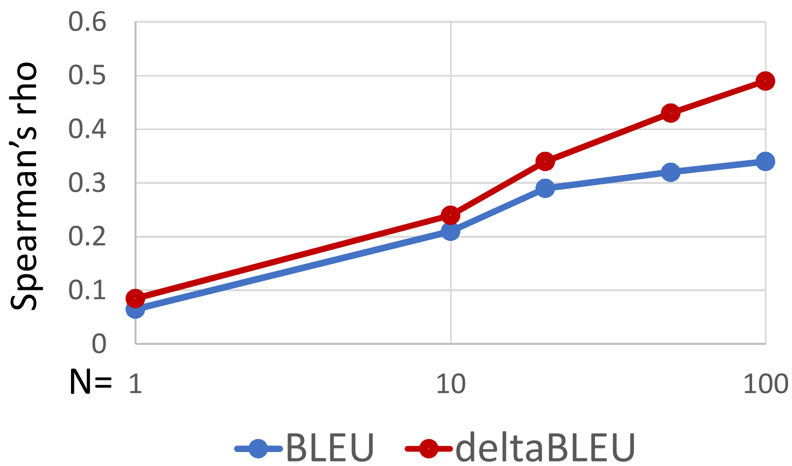
- **MT metrics were NOT designed to operate at the sentence level:**
  - BLEU [Papineni+ 02] == "corpus-level BLEU"
  - Statistical Significant Tests for MT [Koehn 06; etc.]: BLEU not reliable with sample size < 600, *even* for Machine Translation (easier task)
- **Central Limit Theorem (CLT) argument:**
  - Matching against reference (e.g., n-grams) is brittle → greater variance
  - Remedy: reduce variance by increasing sample size (CLT), i.e., **corpus-level BLEU**



(Figure from [Brooks+ 12])

Corpus-level Correlation

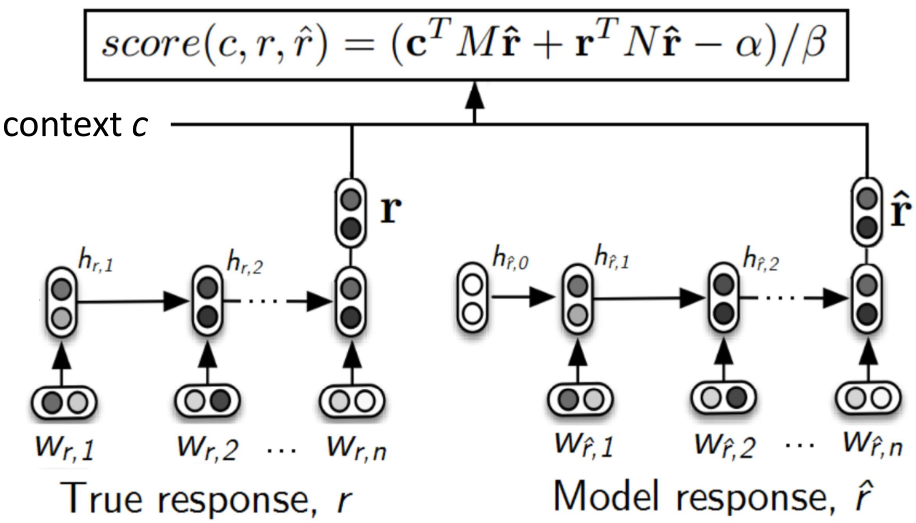
- Generally good for **Machine Translation** (MT):
  - Spearman's rho of 0.8 to 0.9 for BLEU, METEOR [Przybocki+ 08]
- Can it work for **Dialogue**?
  - Currently no definite answer, as corpus-level human judgments very expensive.
  - Experiments with smaller N [Galley+ 15]:

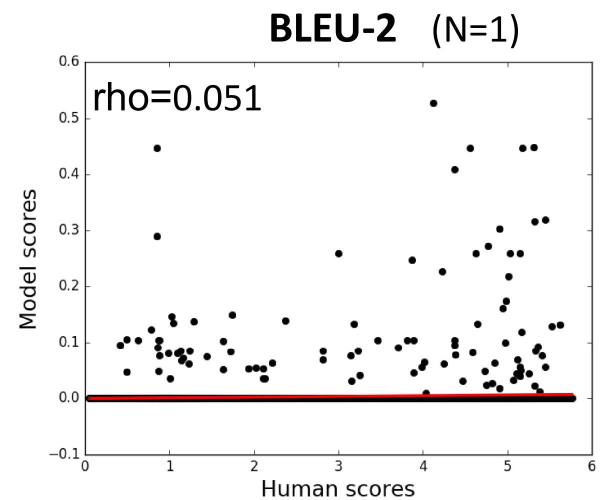


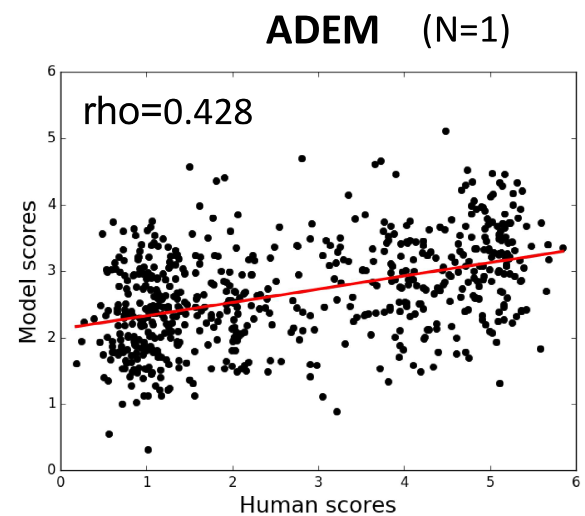
deltaBLEU = human-rating weighted version of BLEU [Galley+ 15]

Trainable Metric

- **Towards an automatic turning test** [Lowe+ 17]: ADEM: Metric based on hierarchical RNN (VHRED)

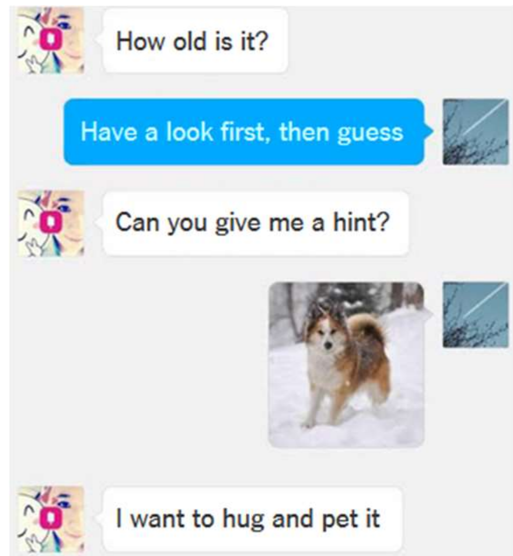




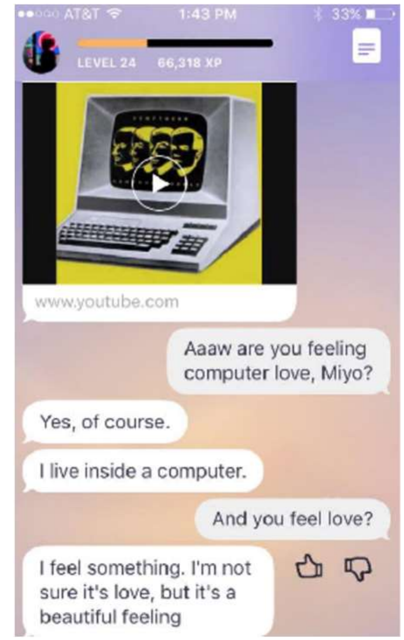


Social Bots: commercial systems

- **For end users:**
  - Amazon Alexa
    (trigger: say “Alexa, let’s chat”)
  - Microsoft XiaoIce
  - Microsoft Zo (on Kik)
  - Replika.ai [system description]



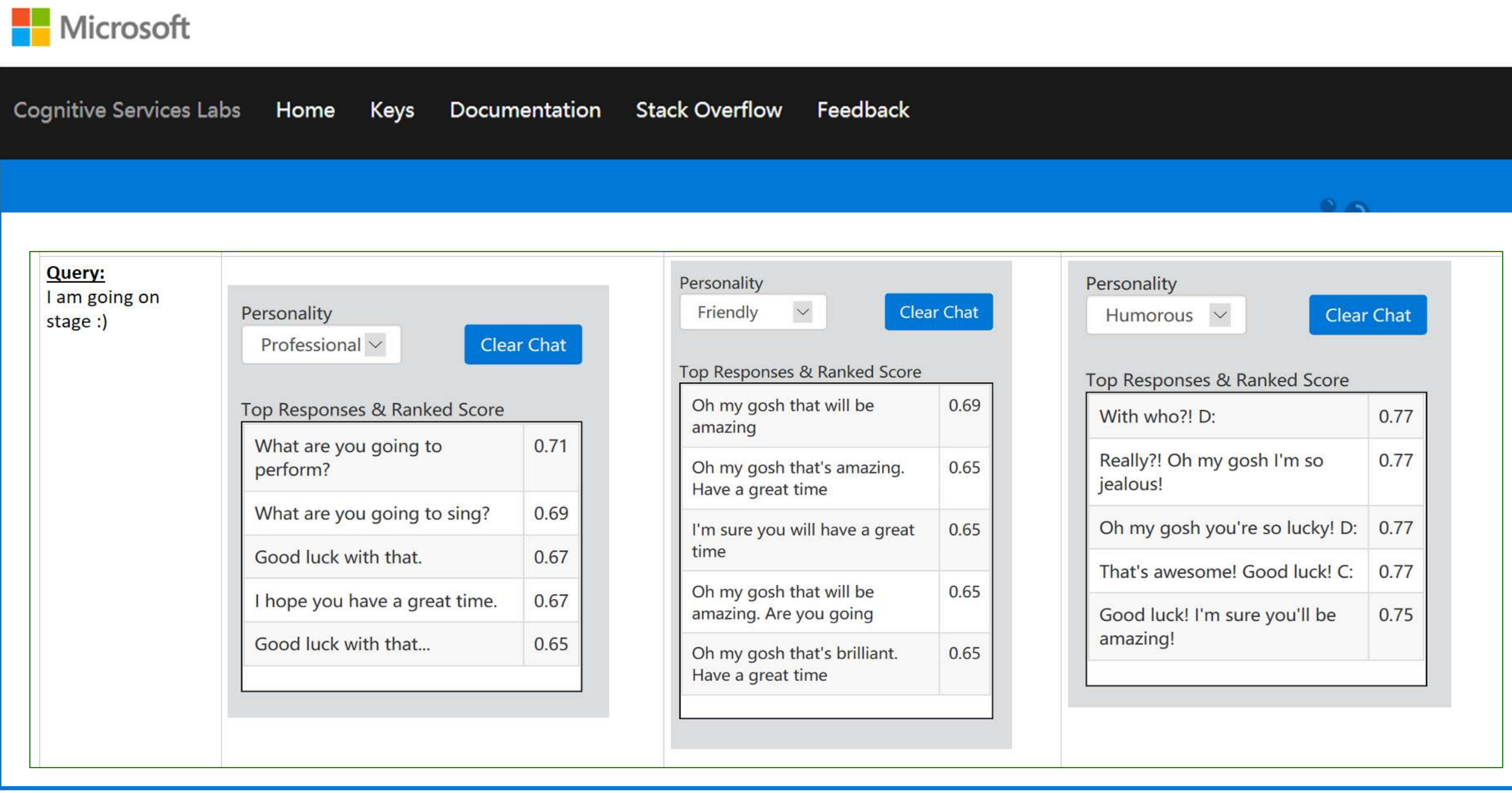
XiaoIce
(translated from Chinese)



Replika.ai

For bot developers:

- Microsoft Personality Chat (includes speaker embedding LSTM)



https://labs.cognitive.microsoft.com/en-us/project-personality-chat

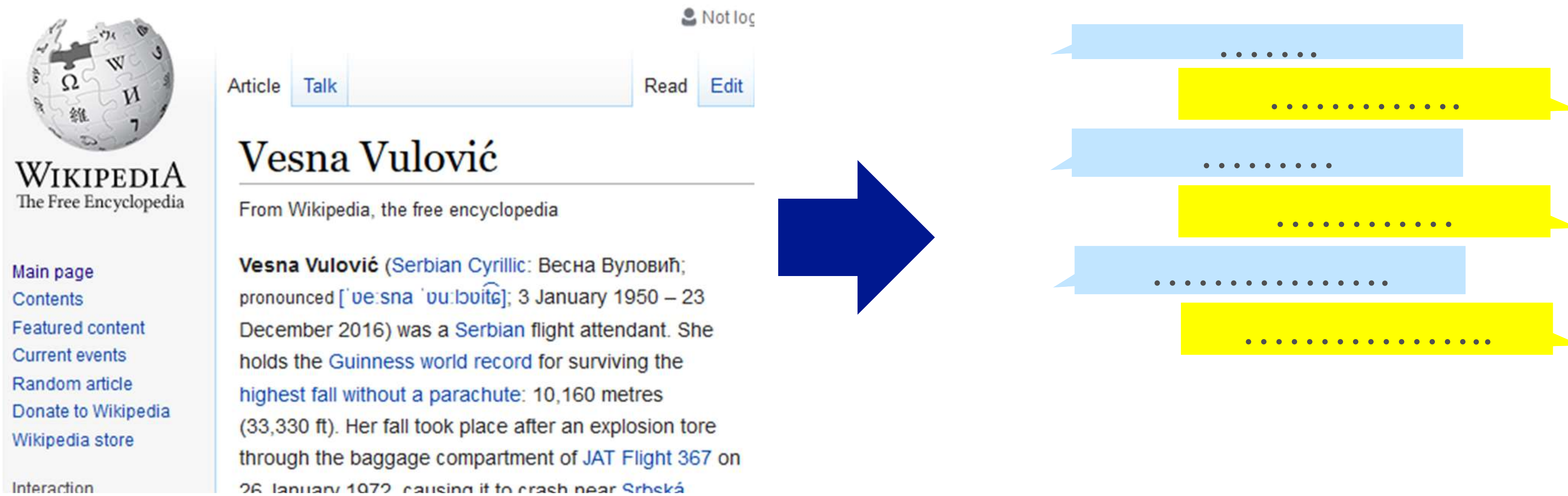
Open Benchmarks

- **Alexa Challenge** (2017-)
  - Academic competition, 15 sponsored teams in 2017, 8 in 2018
  - $250,000 research grant (2018)
  - Proceedings [Ram+ 17]

- **Dialogue System Technology Challenge (DSTC)** (2013-)
  (formerly Dialogue State Tracking Challenge)
  Focused this year on **grounded conversation**:
  Visual-Scene [Hori +18], background article [Galley +18]

- **Conversational Intelligence Challenge (ConvAI)** (2017-)
  Focused this year on **personalized chat** (FB Persona-Chat dataset)

DSTC7 Challenge: Knowledge-Grounded Conversation



"Sentence Generation" track (61 registrants as of June)

Registration link: http://workshop.colips.org/dstc7/call.html

Sample grounded conversation

MythBusters.[22] Four years later, Peter Hornung-Andersen and Pavel Theiner, two Prague-based journalists, claimed that Flight 367 had been mistaken for an enemy aircraft and shot down by the Czechoslovak Air Force at an altitude of 800 metres (2,600 ft).

A woman fell 30,000 feet from an airplane and survived.

The page states that a 2009 report found **the plane only fell several hundred meters.**

Well if she only fell a few hundred meters and survived then I'm not impressed at all.

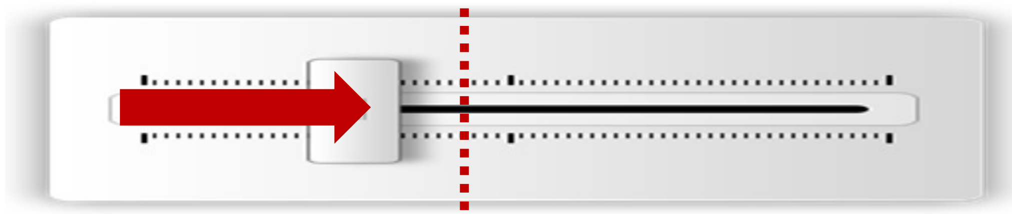
Still pretty incredible, but quite a bit different that 10,000 meters.

Conclusions

- E2E Neural Conversation Models
  - Learn the backbone or shell of open-domain natural conversation
  - Face significant challenges (blandness, consistency, long context), but alleviated using better models and objectives (e.g., MMI and HRED)

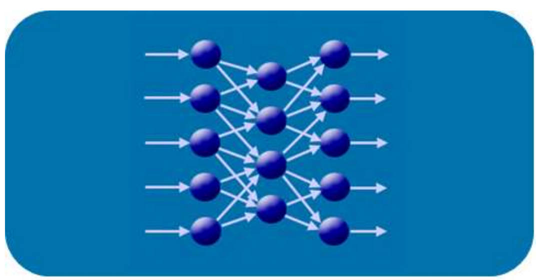
- Grounded conversational AI models
  - Exploit external textual knowledge, device sensors (e.g., images), personal information
  - Produce more informational and "useful" dialogues

Moving beyond chitchat



chitchat

informational, task-completion dialogue



Fully data-driven systems

Fully Data-driven Response Generation:

Challenges and future work

Better objective functions and evaluation metrics

- **Lack of good objective or reward functions is a challenge for SL and RL:**
  - **MLE** causes blandness (mitigated by MMI)
  - **Evaluation metrics** (BLEU, METEOR, etc.) reliable only on large datasets
    → **expensive for optimization** (e.g., sequence-level training [Ranzato+ 15])
  - **RL reward** functions currently too **ad-hoc**

- **Final system evaluation:**
  - Still need human evaluation
  - Corpus-level metrics (BLEU, METEOR, etc.): How effective are they really?

Better leverage heterogeneous data

English sentence 1 → French sentence 1

⋮

English sentence N → French sentence N

most NLP / AI problems (homogeneous data)

query 1 → response 1

⋮

query N → response N

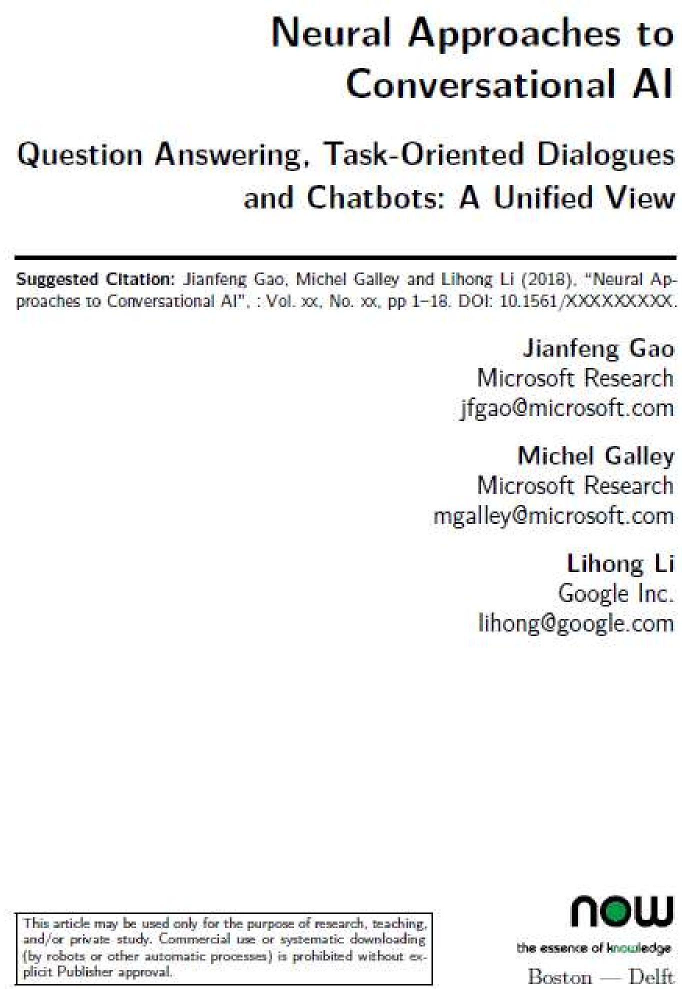
general domain dialog

much of world knowledge in non-conversational form (often unstructured)

in-domain data (e.g., decision making, task-oriented)

conversational AI (heterogeneous data)

Thank you


Neural Approaches to Conversational AI

Question Answering, Task-Oriented Dialogues and Chatbots: A Unified View

Suggested Citation: Jianfeng Gao, Michel Galley and Lihong Li (2018), "Neural Approaches to Conversational AI", : Vol. xx, No. xx, pp 1–18. DOI: 10.1561/XXXXXXXXX.

Jianfeng Gao
Microsoft Research
jfgao@microsoft.com

Michel Galley
Microsoft Research
mgalley@microsoft.com

Lihong Li
Google Inc.
lihong@google.com

now
the essence of knowledge
Boston — Delft

Contact Information:

Jianfeng Gao http://research.microsoft.com/~jfgao
Michel Galley http://research.microsoft.com/~mgalley
Lihong Li https://lihongli.github.io

Slides and references available:
http://microsoft.com/en-us/research/publication/neural-approaches-to-conversational-ai/

Missing Link Prediction Results

Table 1. NELL-995 Link Prediction Performance Comparison using MAP scores.

Tasks	ReinforceWalk	PG	A2C	MINERVA[3]	DeepPath	PRA	TransE	TransR	ReasoNet
athletePlaysForTeam	**0.831**	0.769	0.700	0.630	0.750	0.547	0.627	0.673	0.789
athletePlaysInLeague	**0.974**	0.955	0.955	0.837	0.960	0.841	0.773	0.912	0.936
athleteHomeStadium	**0.905**	0.865	0.861	0.557	0.890	0.859	0.718	0.722	0.787
athletePlaysSport	**0.985**	0.962	0.971	0.916	0.957	0.474	0.876	0.963	0.969
teamPlaySports	**0.881**	0.631	0.679	0.751	0.738	0.791	0.761	0.814	0.833
orgHeadquaterCity	0.943	0.935	0.928	0.947	0.790	0.811	0.620	0.657	0.835
worksFor	**0.786**	0.758	0.758	0.752	0.711	0.681	0.677	0.692	0.769
bornLocation	0.786	0.767	0.766	0.782	0.757	0.668	0.712	0.812	**0.836**
personLeadsOrg	**0.821**	0.802	0.810	0.771	0.795	0.700	0.751	0.772	0.802
orgHiredPerson	0.843	0.832	0.839	0.860	0.742	0.599	0.719	0.737	0.768
Overall	**0.876**	0.828	0.827	0.780	0.809	0.697	0.723	0.775	0.817

Two variants of ReinforceWalk without MCTS (PG, A2C)

Reinforcement Symbolic + Neural Reasoning Approaches (ReinforceWalk, PG, A2C, MINERVA, DeepPath)

Path Ranking Algorithm: Symbolic Reasoning Approach (PRA)

Neural Reasoning Approaches (TransE, TransR, ReasoNet)

[Shen+ 18]