# Neural Approaches to Machine Reading Comprehension and Dialogue

**Jianfeng Gao**

Thanks for slides by **Bill Dolan, Lihong Li, Xiujun Li, Yelong Shen**

Microsoft AI & Research

September 18, 2017, UC Berkeley

## Contact Information:

www.microsoft.com/en-us/research/people/jfgao/

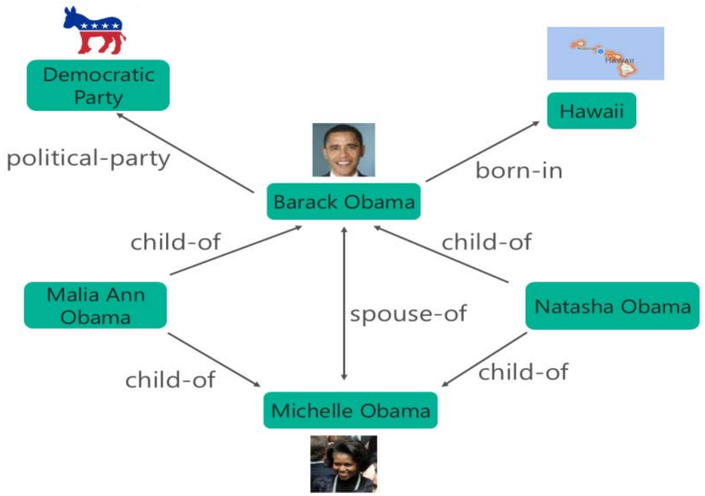
## Collaborators:

Faisal Ahmed, Asli Celikyilmaz, Ming-Wei Chang, Weizhu Chen, Yun-Nung Chen, Li Deng, Bhuwan Dhingra, Bill Dolan, Michel Galley, Po-Sen Huang, Sungjin Lee, Lihong Li, Xiujun Li, Zachary Lipton, Baolin Peng, Yelong Shen, Alessandro Sordoni, et al.

# 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: production system

https://en.wikipedia.org/wiki/Production_system_(computer_science)

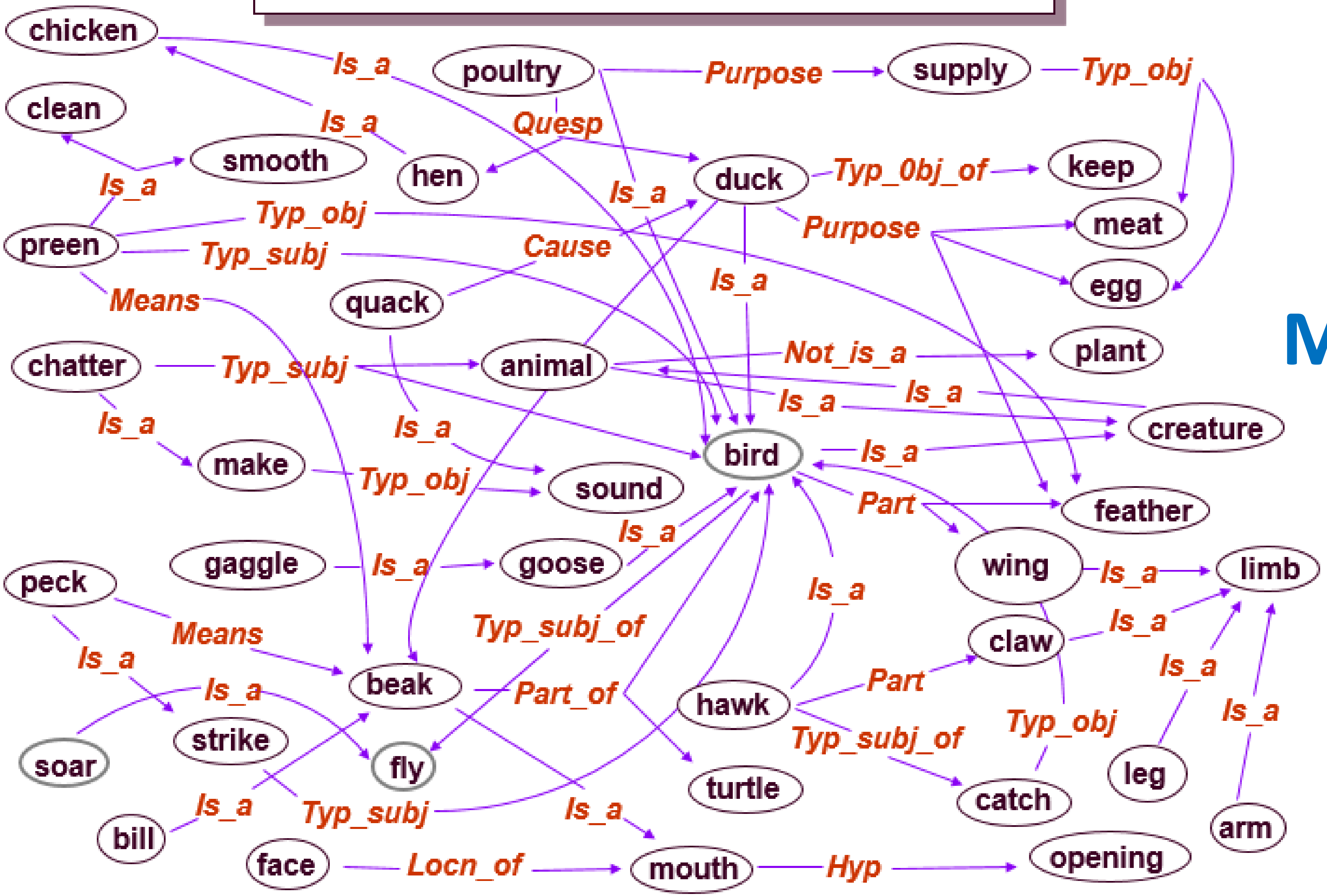
- Production rules
  - condition—action pairs
  - Represent (world) knowledge as a graph
- Working memory
  - Contains a description of the current state of the world in a reasoning process
- Recognizer-act controller
  - Update working memory by searching and firing a production rule

- A case study: MSR MindNet [Dolan+ 93; Richardson+ 98]

# Case study of Question Answering with MindNet

- Build a MindNet graph from:
  - Text of dictionaries
  - Target corpus, e.g. an encyclopedia (Encarta 98)
- Build a dependency graph from query
- Model QA as a graph matching procedure
  - Heuristic fuzzy matching for synonyms, named entities, wh-words, etc.
  - Some common sense reasoning (e.g. dates, math)
- Generate answer string from matched subgraph
  - Including well-formed answers that didn’t occur in original corpus

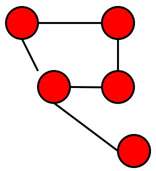
## *Fragment of lexical space surrounding “bird”*

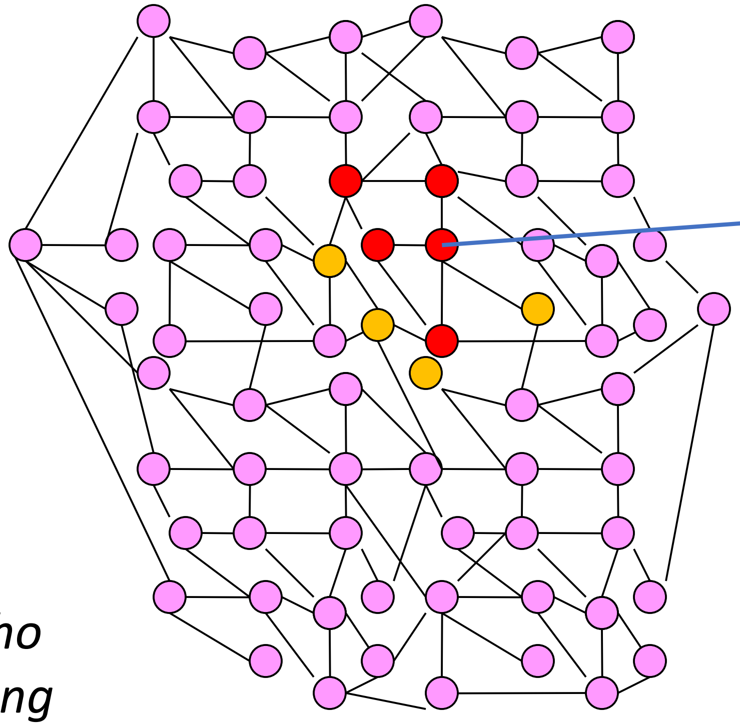


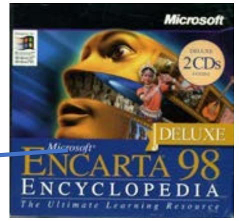

# MindNet

# Fuzzy Match against MindNet

Input LF:





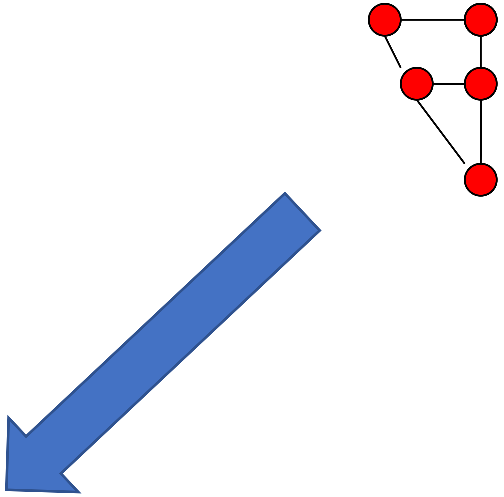



_Lincoln, Abraham_

*Who assassinated Abraham Lincoln?*

*American actor John Wilkes Booth, who was a violent backer of the South during the Civil War, shot Abraham Lincoln at Ford's Theater in Washington, D.C., on April 14, 1865.*

# Generate output string



*“John Wilkes Booth shot Abraham Lincoln”*

# Worked beautifully!

- Just not very often…
- What went wrong?
  - One major reason: paraphrase alternations

"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).

## Symbolic Space

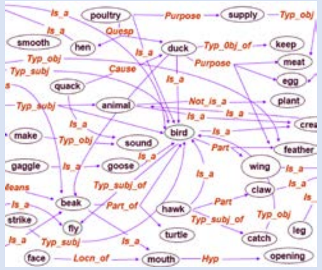
- **Knowledge Representation**
  - *Explicitly* store a BIG but incomplete knowledge graph (KG)
  - Words, relations, templates
  - High-dim, discrete, sparse vectors
- **Inference**
  - Slow on a big KG w. millions of paths
  - Keyword/template matching is sensitive to paraphrase alternations
- **Human 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:

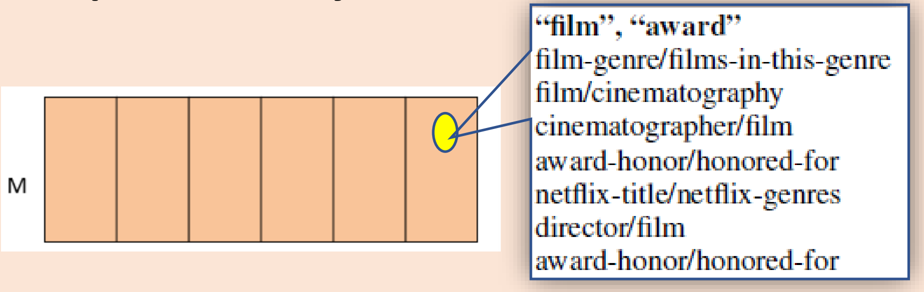
*"Fifteen men on the dead man's chest-Yo-ho-ho, and a bottle of rum!"* in the high, old tottering voice that seemed to have been tuned and broken at the capstan bars. Then he rapped on the door with a bit of stick like a handspike that he carried, and when my father appeared, called roughly for a glass 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

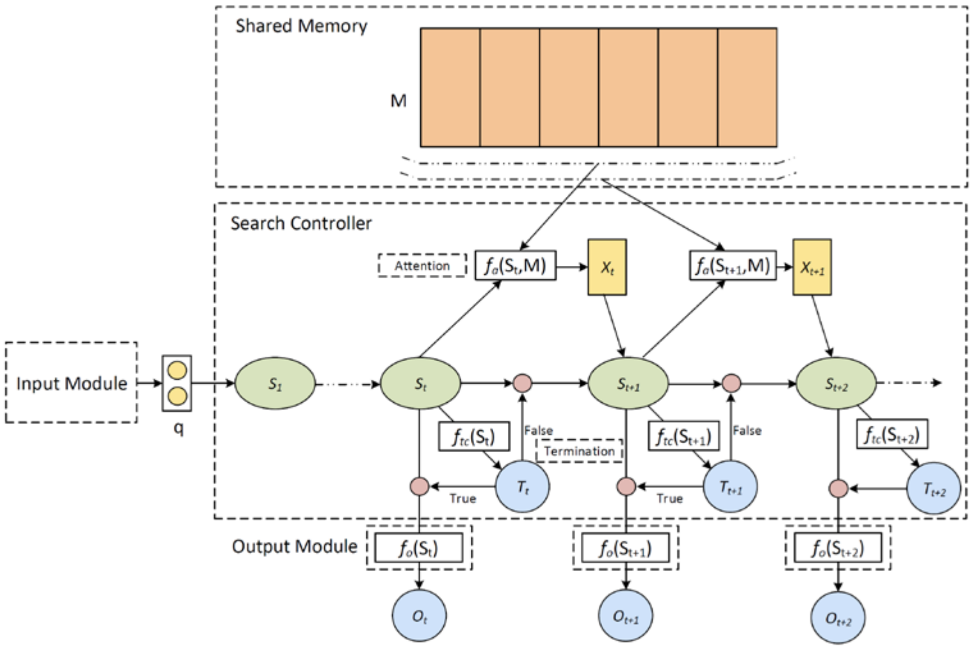



## Neural Space

- **Knowledge Representation**
  - *Implicitly* store entities and structure of KG in a *compact* way that is more generalizable
  - Semantic concepts/classes
  - Low-dim, cont., dense vectors shaped by KG
- **Inference**
  - Fast on compact memory
  - Semantic matching is robust to paraphrase alternations
- **Computationally efficient but not human comprehensible *yet***



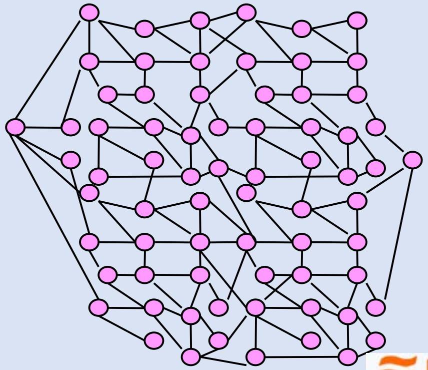
# ReasoNet with Shared Memory




- **Input/output modules** are task-specific
- **Shared memory** encodes task-specific knowledge
- **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
- *Shared memory and search controller are jointly learned*

[Shen+ 16a]

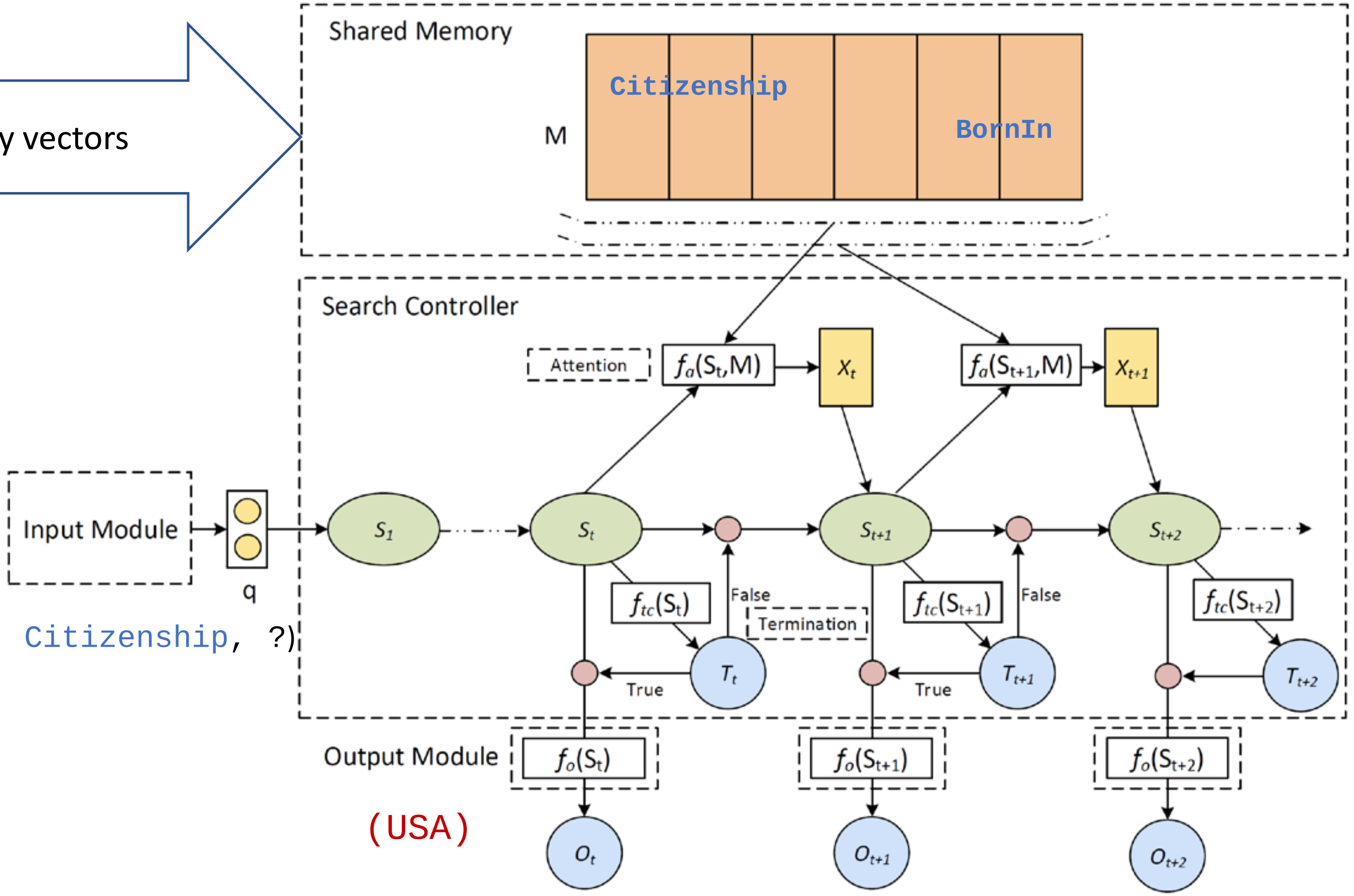
# Joint learning of Shared Memory and Search Controller



Embed KG to memory vectors

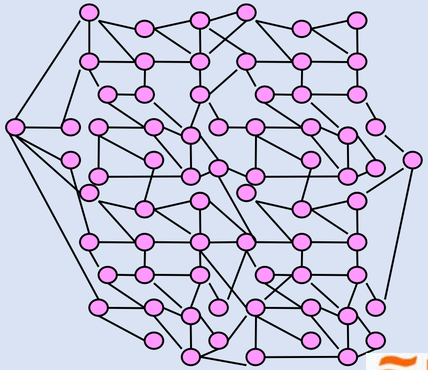
Training samples from KG:
(Obama, BornIn, Hawaii)
(Hawaii, PartOf, USA)
…
(*h*, *r*, *t*)
…
(Obama, Citizenship,?)->(USA)

(Obama, Citizenship, ?)

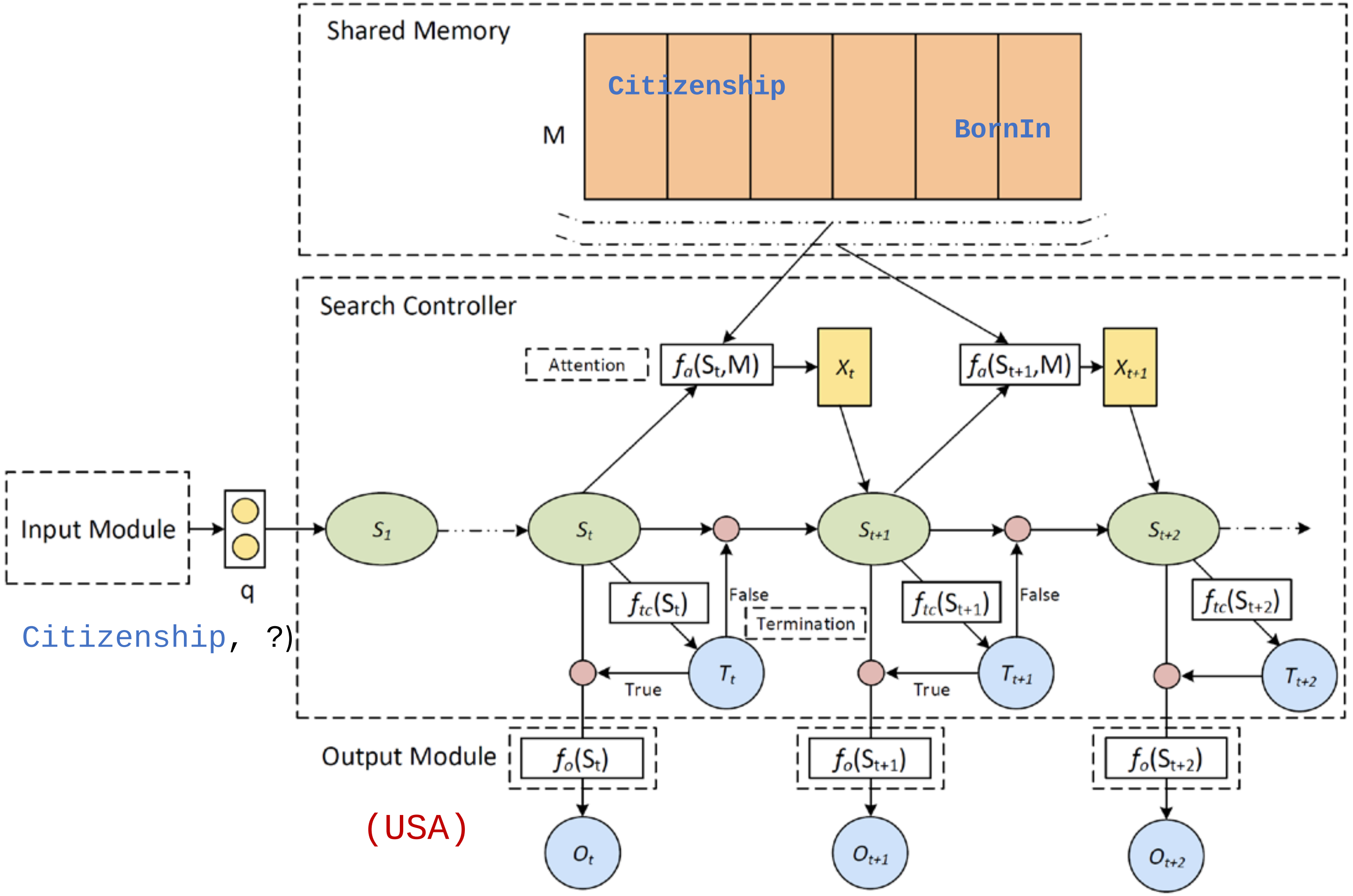



(USA)

# Joint learning of Shared Memory and Search Controller

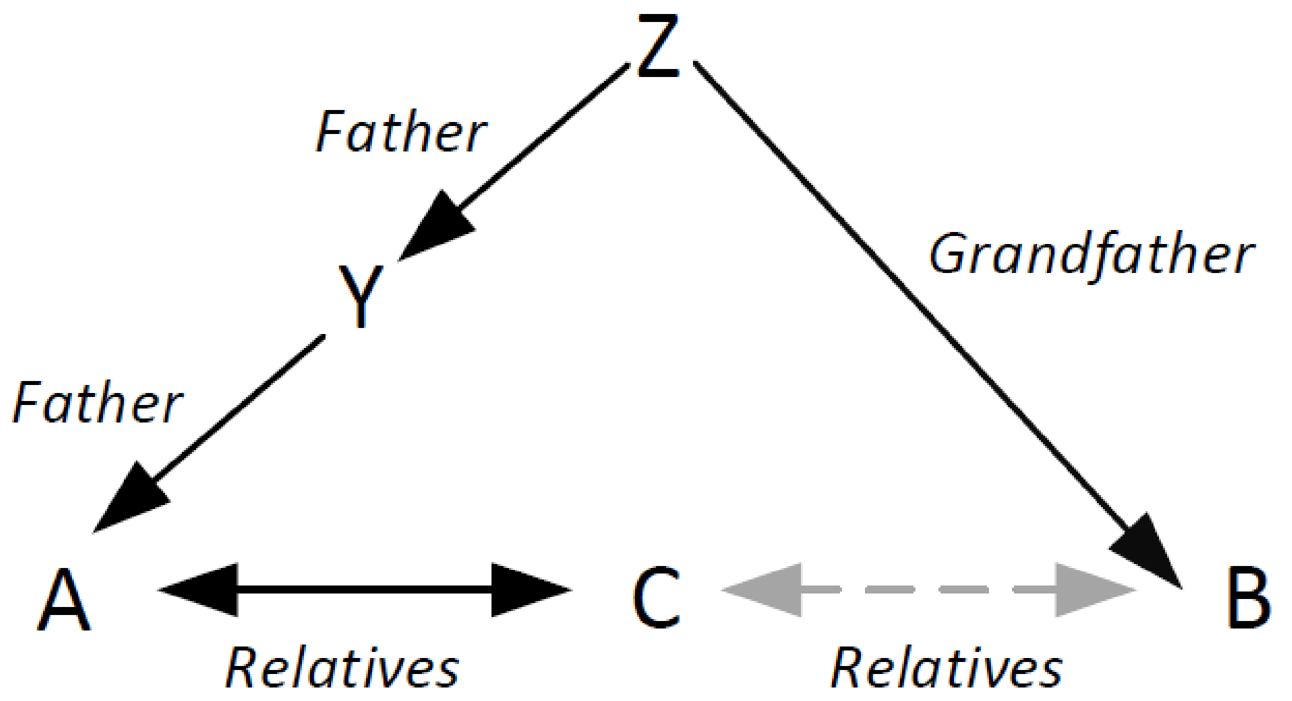


Training samples from KG:
(Obama, BornIn, Hawaii)
(Hawaii, PartOf, USA)
…
(*h*, *r*, *t*)
…
(Obama, Citizenship,?)->(USA)




(Obama, Citizenship, ?)

(USA)



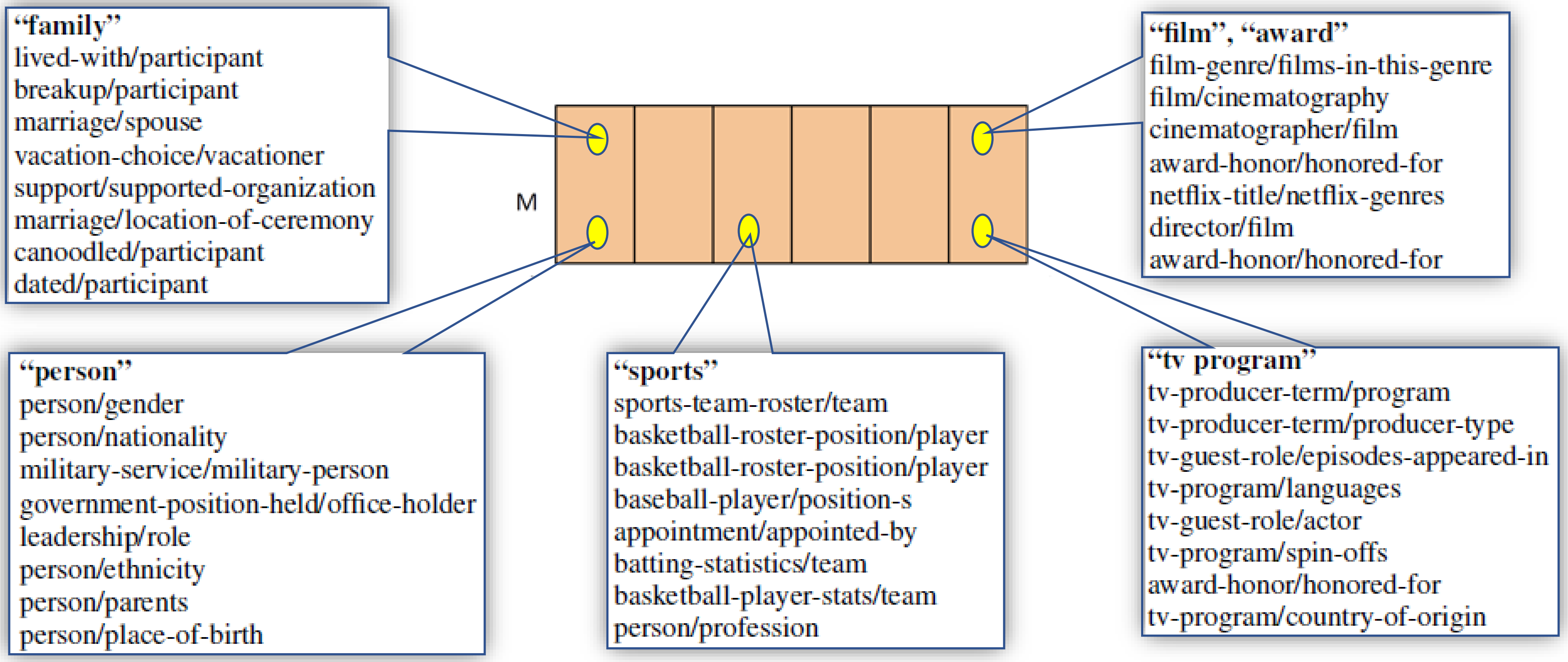

| Space | Inference Path |
|---|---|
| Symbolic | $B \xrightarrow{\text{Grandfather}^{-1}} Z \xrightarrow{\text{Father}} Y \xrightarrow{\text{Father}} A \xrightarrow{\text{Relatives}} C$ |
| Neural | $B \xrightarrow{\text{Relatives}} ? \Rightarrow A \xrightarrow{\text{Relatives}} C$ |

# The Knowledge Base Question Answering Results on WN18 and FB15K

| Model | Additional Information | WN18 | | FB15k | |
|---|---|---|---|---|---|
| | | Hits@10 (%) | MR | Hits@10 (%) | MR |
| SE (Bordes et al., 2011) | NO | 80.5 | 985 | 39.8 | 162 |
| Unstructured (Bordes et al., 2014) | NO | 38.2 | 304 | 6.3 | 979 |
| TransE (Bordes et al., 2013) | NO | 89.2 | 251 | 47.1 | 125 |
| TransH (Wang et al., 2014) | NO | 86.7 | 303 | 64.4 | 87 |
| TransR (Lin et al., 2015b) | NO | 92.0 | 225 | 68.7 | 77 |
| CTransR (Lin et al., 2015b) | NO | 92.3 | 218 | 70.2 | 75 |
| KG2E (He et al., 2015) | NO | 93.2 | 348 | 74.0 | 59 |
| TransD (Ji et al., 2015) | NO | 92.2 | 212 | 77.3 | 91 |
| TATEC (García-Durán et al., 2015) | NO | - | - | 76.7 | 58 |
| NTN (Socher et al., 2013) | NO | 66.1 | - | 41.4 | - |
| DISTMULT (Yang et al., 2014) | NO | 94.2 | - | 57.7 | - |
| STransE (Nguyen et al., 2016) | NO | 94.7 (93) | 244 (**206**) | 79.7 | 69 |
| RTransE (García-Durán et al., 2015) | Path | - | - | 76.2 | 50 |
| PTransE (Lin et al., 2015a) | Path | - | - | 84.6 | 58 |
| NLFeat (Toutanova et al., 2015) | Node + Link Features | 94.3 | - | 87.0 | - |
| Random Walk (Wei et al., 2016) | Path | 94.8 | - | 74.7 | - |
| **ReasoNet (Shen+ 16a)** | NO | **95.3** | 249 | **92.7** | **38** |

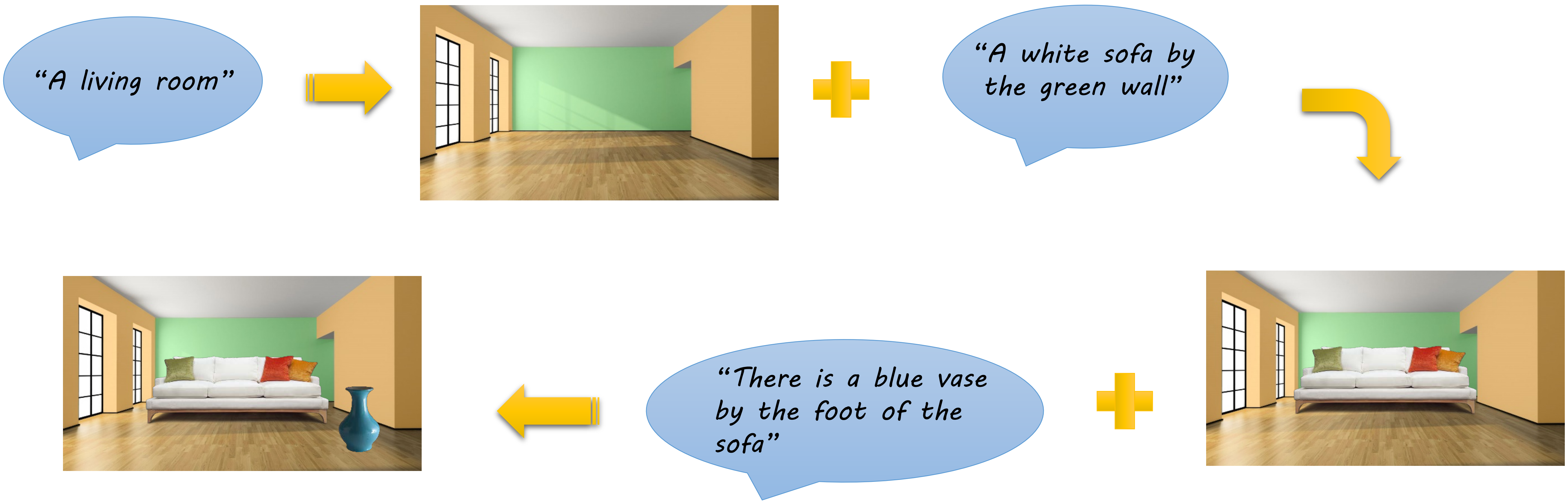
# Shared Memory: long-term memory to store learned knowledge, like human brain

- 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.



# Visualization of Reasoning in MRC Models

- **Translate Natural Language to Image**
- **Multi-step Image Editing via Dialogue**



# Dialogue

- What kinds of Problems?
- Deep RL for dialogues
- Three case studies

# Aspirational Goal: Enterprise Assistant

*Task Completion*

*Decision Support*

*Info Consumption*

*Task Completion*

Where are sales lagging behind our forecast?

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

Do you know why?

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]?

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

Thanks

# Three types of dialogue systems

- Social bot (not to be discussed today)
  - Microsoft XiaoIce,
  - MSR Data-Driven Conversation
- Task-completion bot
  - Movie ticket booking
  - Hotels booking
  - Travel assistant
- Info bot
  - Find the closest Starbucks with drive-thru
  - Find a family-friendly movie directed by Andrew Stanton near Redmond for upcoming weekend afternoons

Goal-oriented (slot-filling) dialogues

# MSR’s 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: | wan na talk about it ? i 'm here if you wan na talk x |
| H: | well , i 'm just tired |
| M: | me too , i 'm here if you wan na 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 2 pm work for you ? |
| M: | works for me . |
| H: | well let ‘s say 2 pm then i ‘ll see you there |
| M: | sounds good . |

- Generating responses vs. retrieving responses
- Easy to incorporate contextual info via embedding
    - User profile – personalized conversation
    - knowledge bases – grounded conversation
- The engine is E2E learned from conversation experience
    - Learning a goal-oriented conversation engine via RL
- **Ongoing: knowledge-grounded dialogue**

[Sordoni+ 15; Li+ 16a]

# Conversation in the real world

- **conversation history**
- **persona**
- **user data (bio, social graph, etc.)**
- **visual signal (camera, picture etc.)**
- **knowledge base**
- mood
- geolocation
- time

# Image-grounded dialogue

Neural approaches allow language models to be grounded in the world, i.e., link language to real-world signals such as images, machine state, sensor data from biomedical devices.


Output of a neural conversation model trained on 250K Twitter conversations sparked by a tweeted photo

[Mostafazadeh+ 17]

# An example dialogue of MovieBot

```
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?
Turn 8 usr: 10:00 pm
2 tickets for you
cific place 11 theater
```

Some of our dialogues can be more complex:

- Natural language understanding errors
  → reason under uncertainty
- Constraint violation
  → revise information collected earlier

Source code available on https://github.com/MiuLab/TC-Bot

# Slot-filling dialogues

- Slot: information to be filled in before completing a task
  - For movie-bot: movie-name, theater, number-of-tickets, price, …
- Dialog act (intent)
  - Inspired by speech act theory (communication as action)

    request, confirm, inform, thank-you, …
  - Some may take parameters:

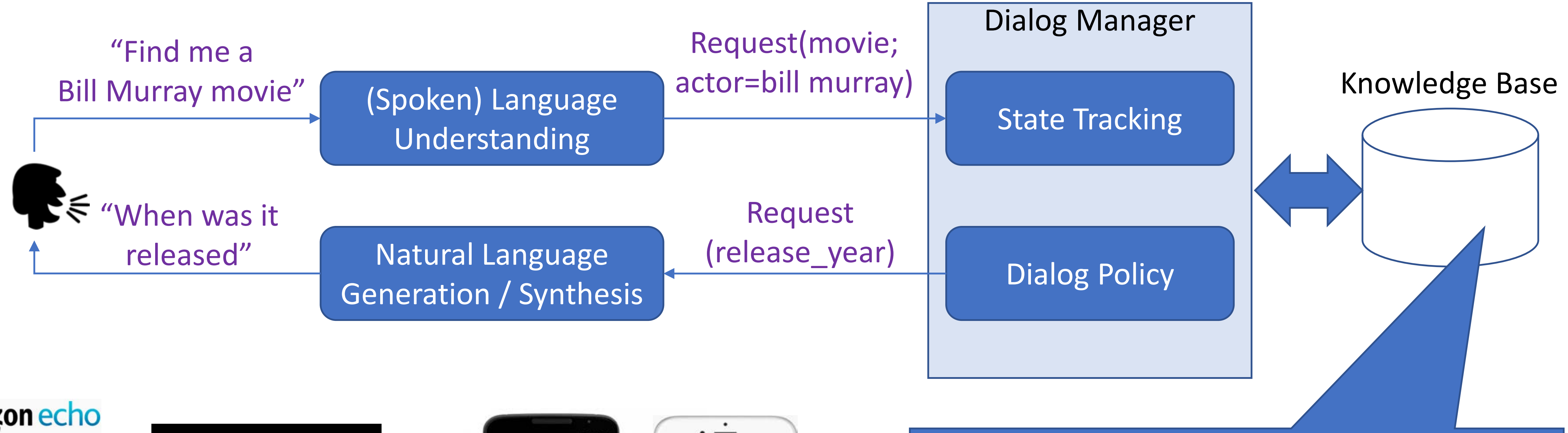
    request(price)

    confirm(moviename="kungfu panda")

    inform(price=$10)

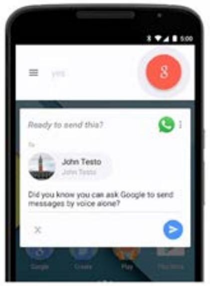
    thank-you()

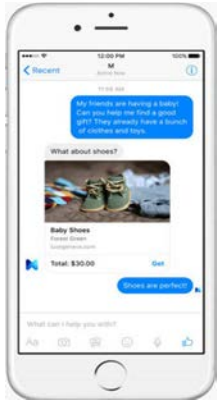
# Multi-turn (goal-oriented) dialogue







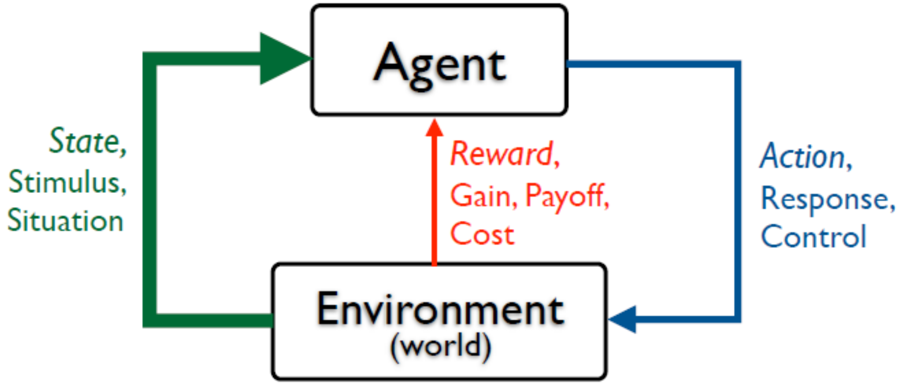


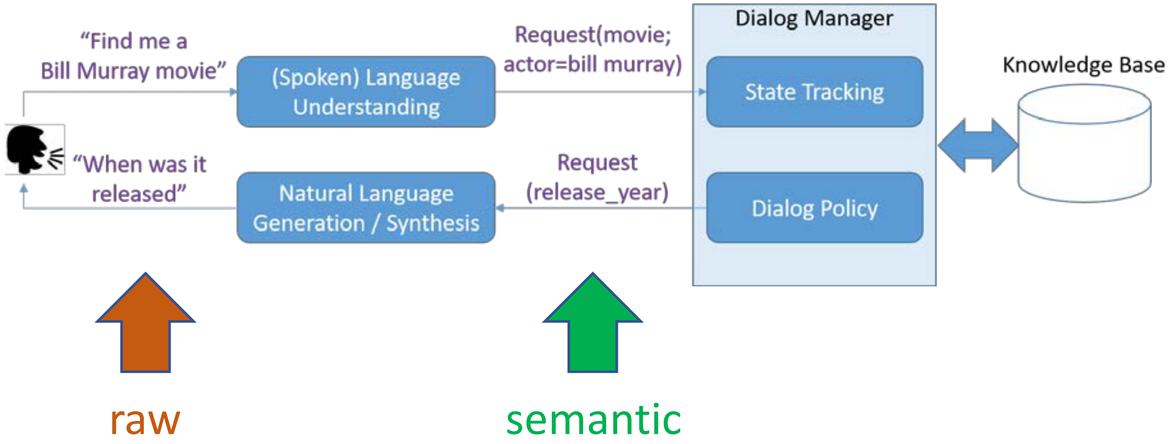


**Entity-Centric Knowledge Base**

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

# Conversation as RL

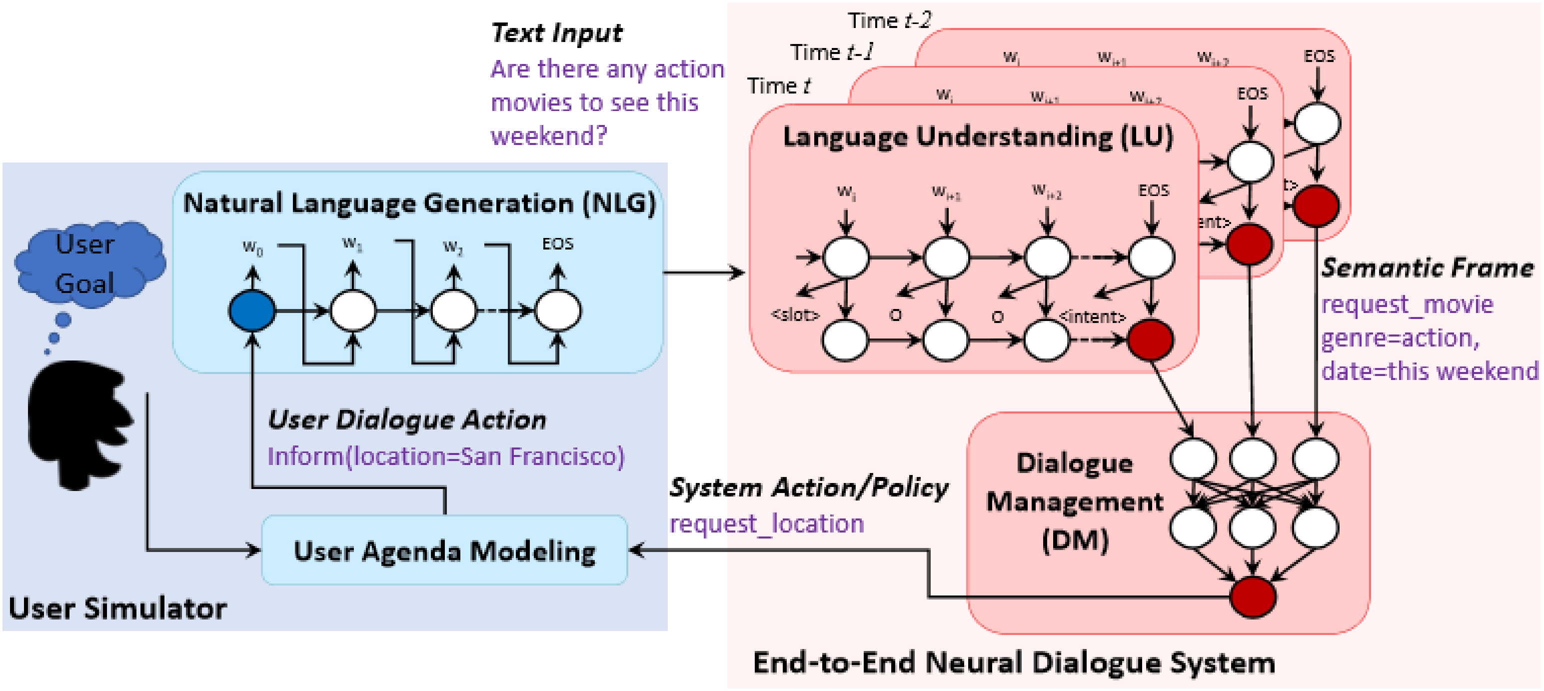



- Observation / Action
  - Raw utterance (natural language form)
  - Semantic representation (dialog-acts)
- Reward
  - $+10$ upon termination if succeeded
  - $-10$ upon termination if failed
  - $-1$ per turn
- State
  - Explicitly defined (POMDP-based, …)
  - Implicitly defined (RNNs)

Earlier examples: [Levin+ 00; Singh+ 02; Williams & Young 07]

# A user simulator for RL and evaluation




- Robustness: automatic action selection based on uncertainty by RL
- Flexibility: allow user-initiated behaviors
- Reproducibility: a R&D setting that allows consistent comparisons of competing methods

[Li+ 17] https://github.com/MiuLab/TC-Bot

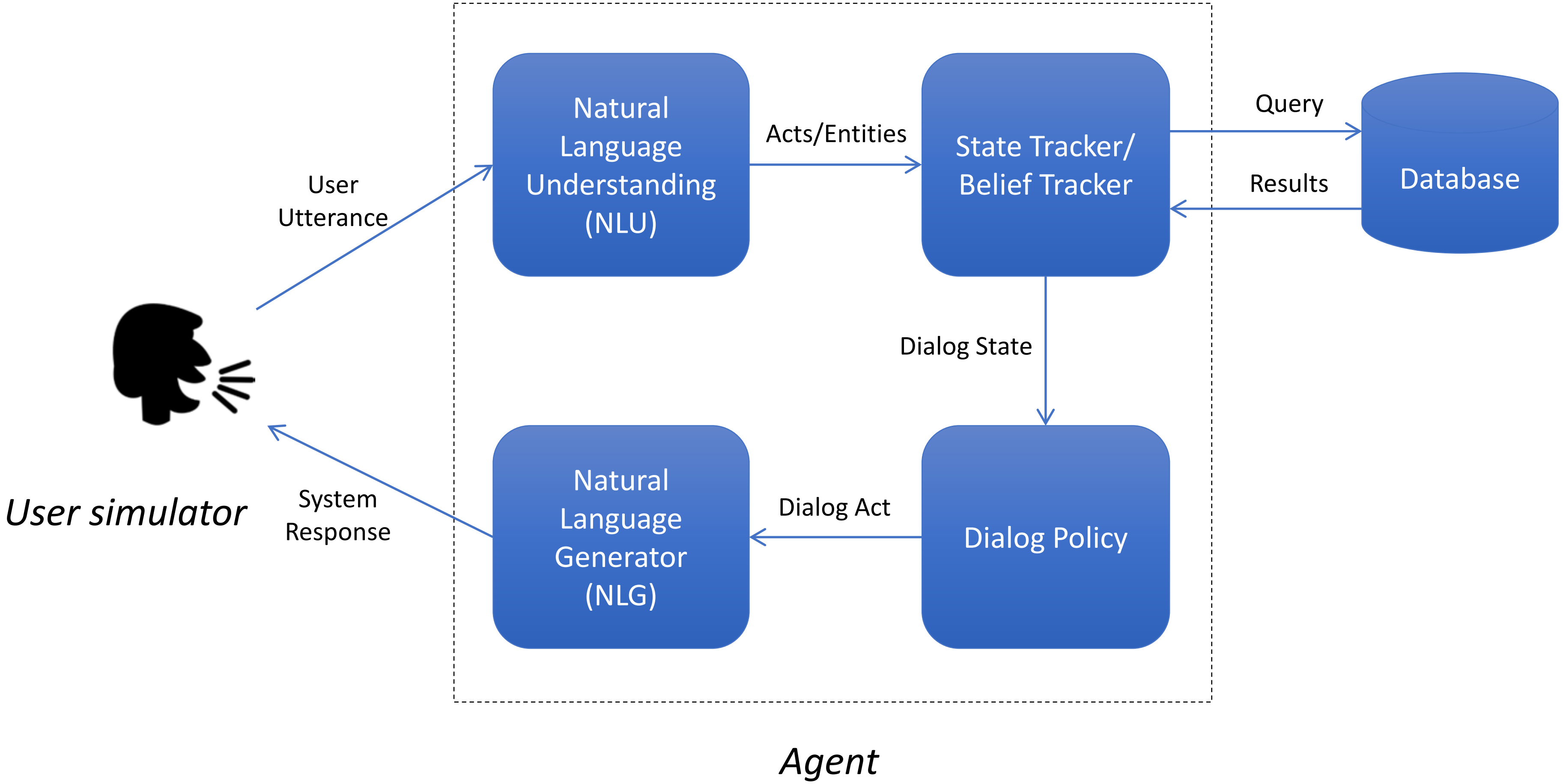
# Three case studies

- Info bot: end-to-end training with non-differentiable knowledge base [Dhuwan+ 17]
- Composite task completion bot with Hierarchical RL [Peng+ 17]
- Task-completion bot: efficient exploration for domain extension [Zachary+ 17]

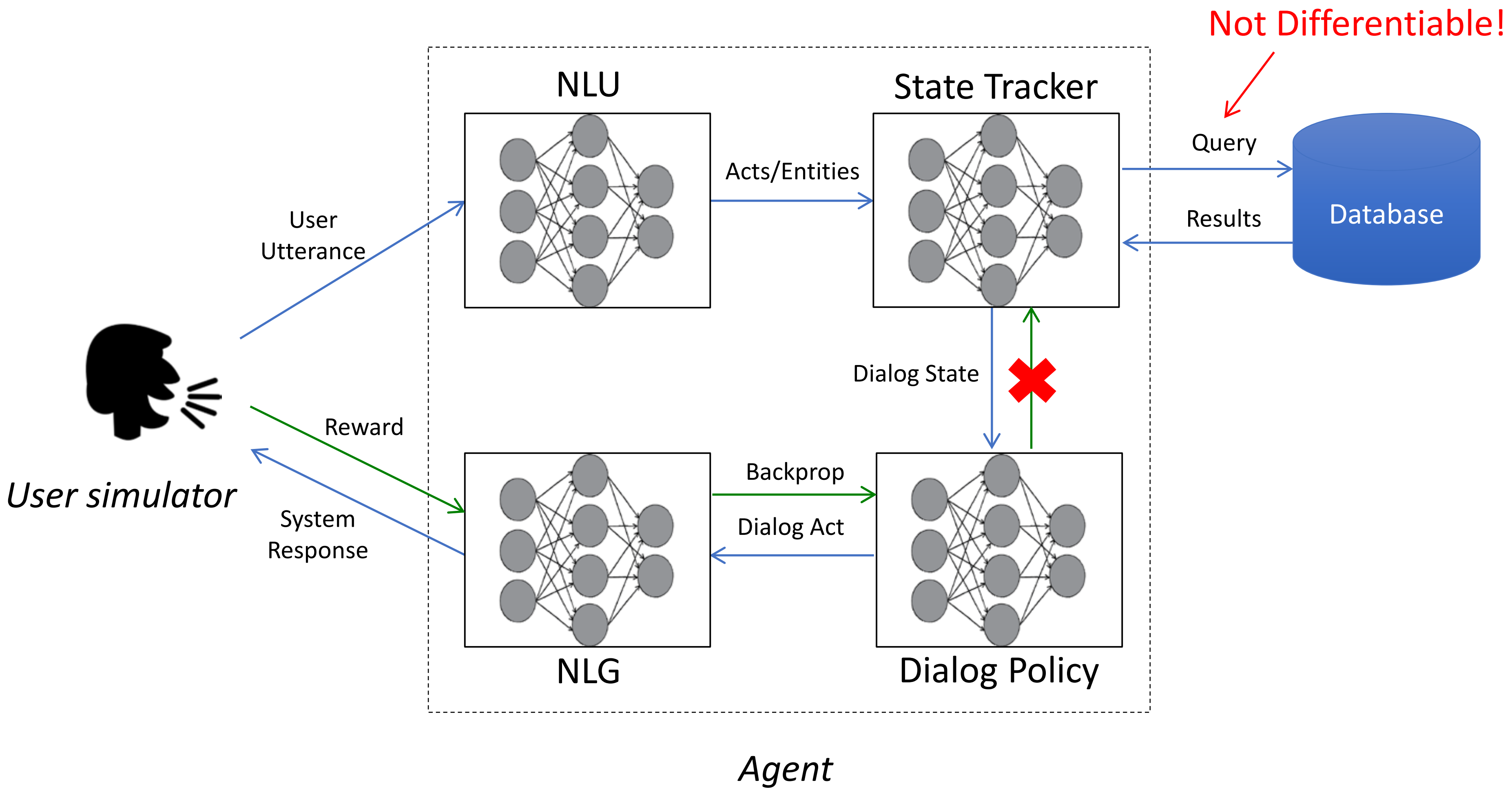
# InfoBot as an interactive search engine

- Problem setting
  - User is looking for a piece of information from one or more tables/KBs
  - System must iteratively ask for user constraints (“slots”) to retrieve the answer
- A general rule-based approach
  - Given current beliefs, ask for slot with maximum uncertainty
  - Works well in most cases but,
    - Has no notion of what the user is likely to be looking for or likely to know
    - No principled way to deal with errors/uncertainty in language understanding

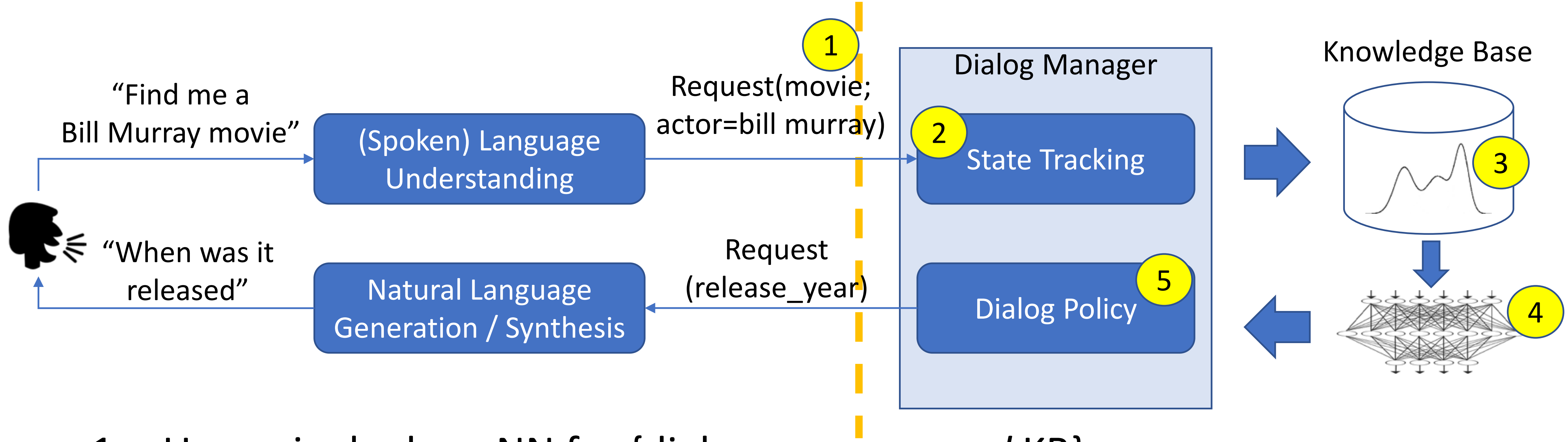
# InfoBot as an interactive search engine



# Deep Reinforcement Learning



# Our end-to-end approach




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 attention for 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:

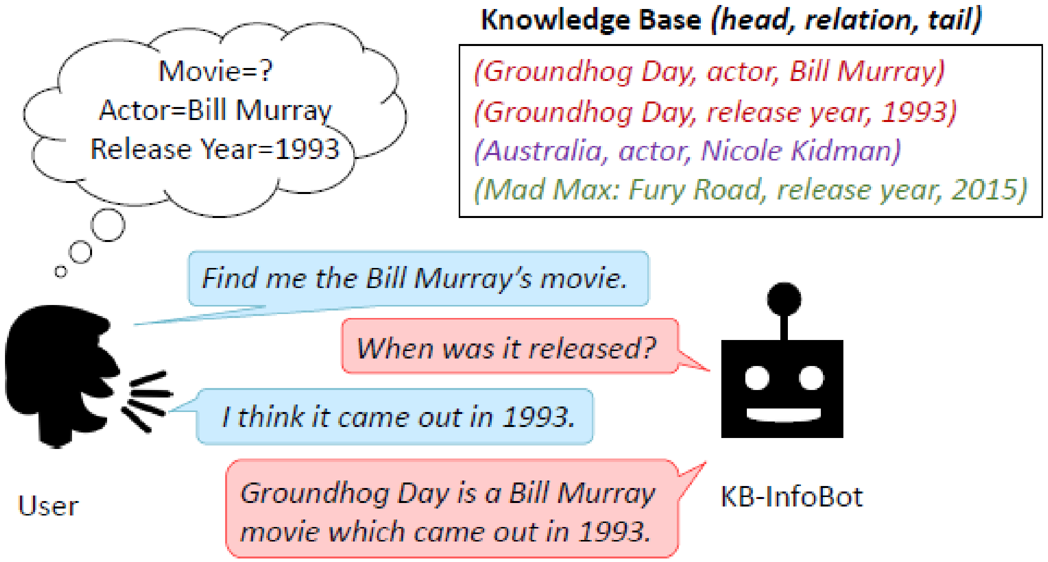
$$\Pr(\text{"GroundhogDay"}) \propto \Pr(\text{Actor} = \text{"Bill Murray"}) \cdot \Pr(\text{ReleaseYear} = \text{"1993"}) \cdots$$

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

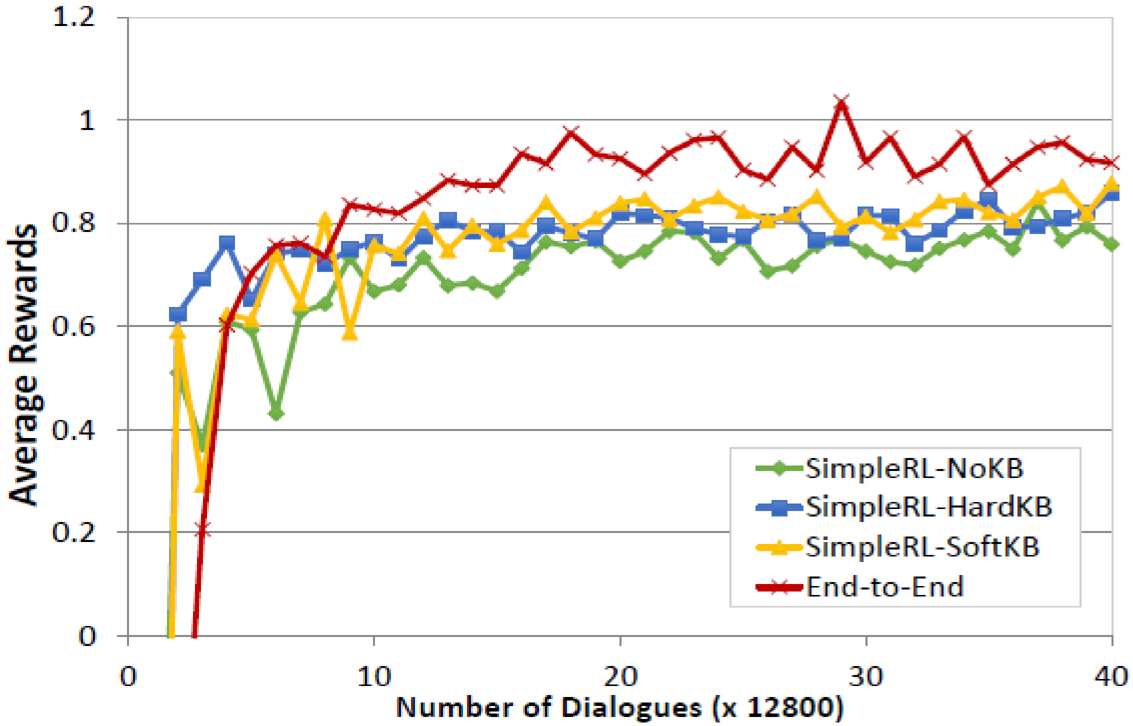
- Soft KB-lookup: sample a movie according to the posterior
  - Randomization results in differentiability (similar to policy gradient alg.)
  - As opposed to using SQL queries to look up results deterministically

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

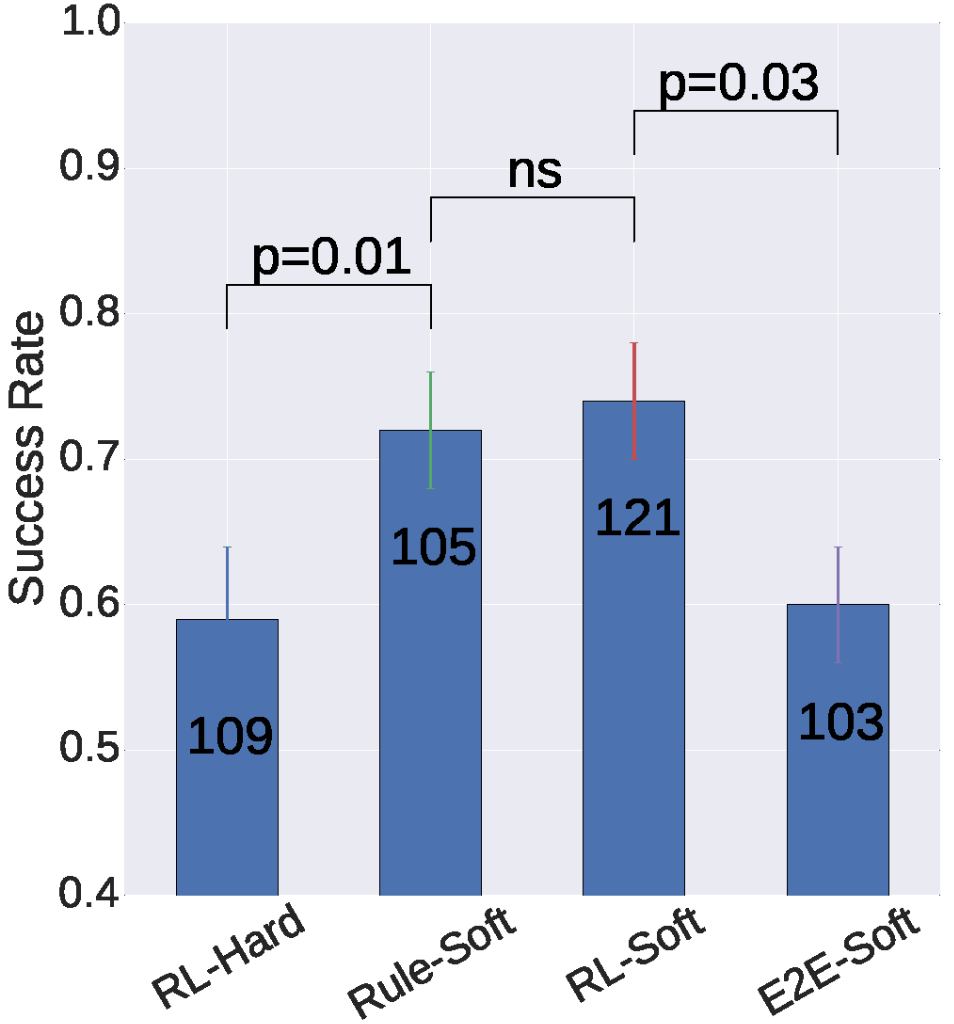
# Result on IMDB using KB-InfoBot w/ simulated users

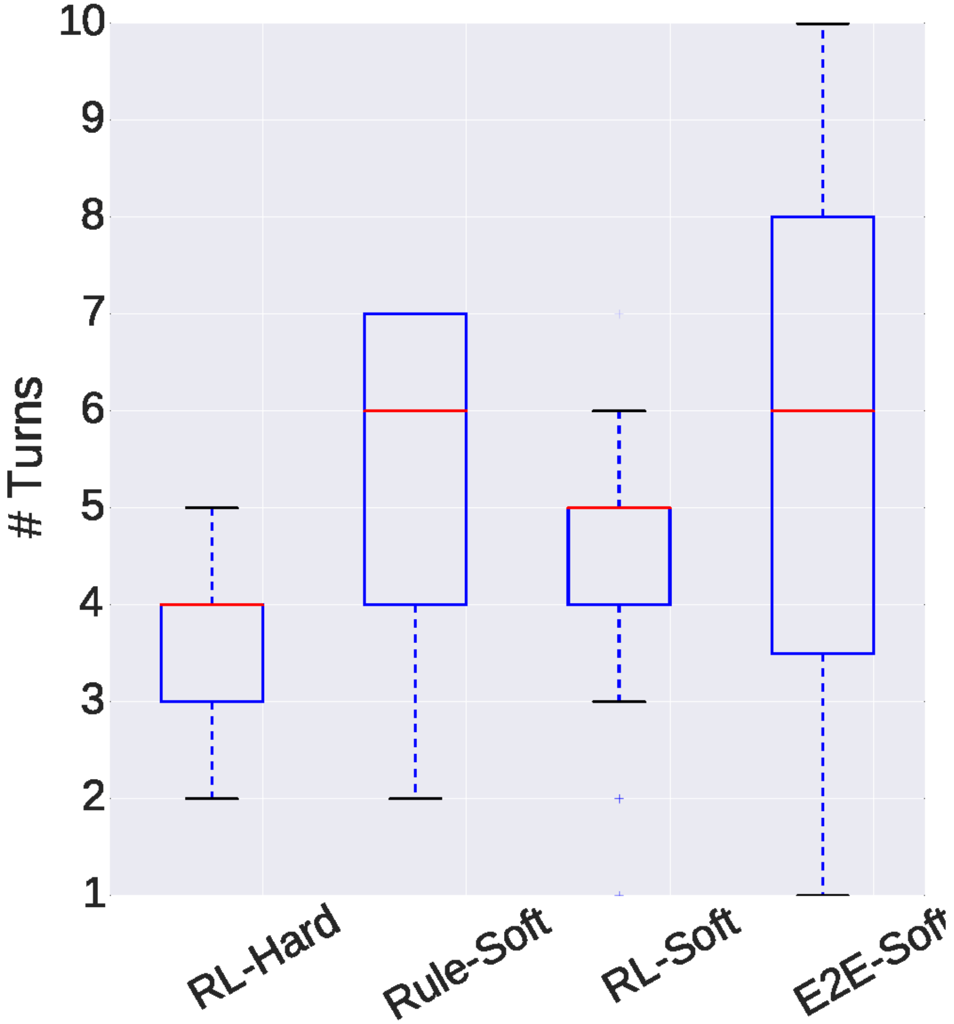



| Agent | Success Rate | Avg Turns | Avg Reward |
|---|---|---|---|
| Rule-Soft | 0.76 | 3.94 | 0.83 |
| RL-Hard | 0.75 | **3.07** | 0.86 |
| RL-Soft | 0.80 | 3.37 | 0.98 |
| E2E-RL | **0.83** | 3.27 | **1.10** |



# Results on real users



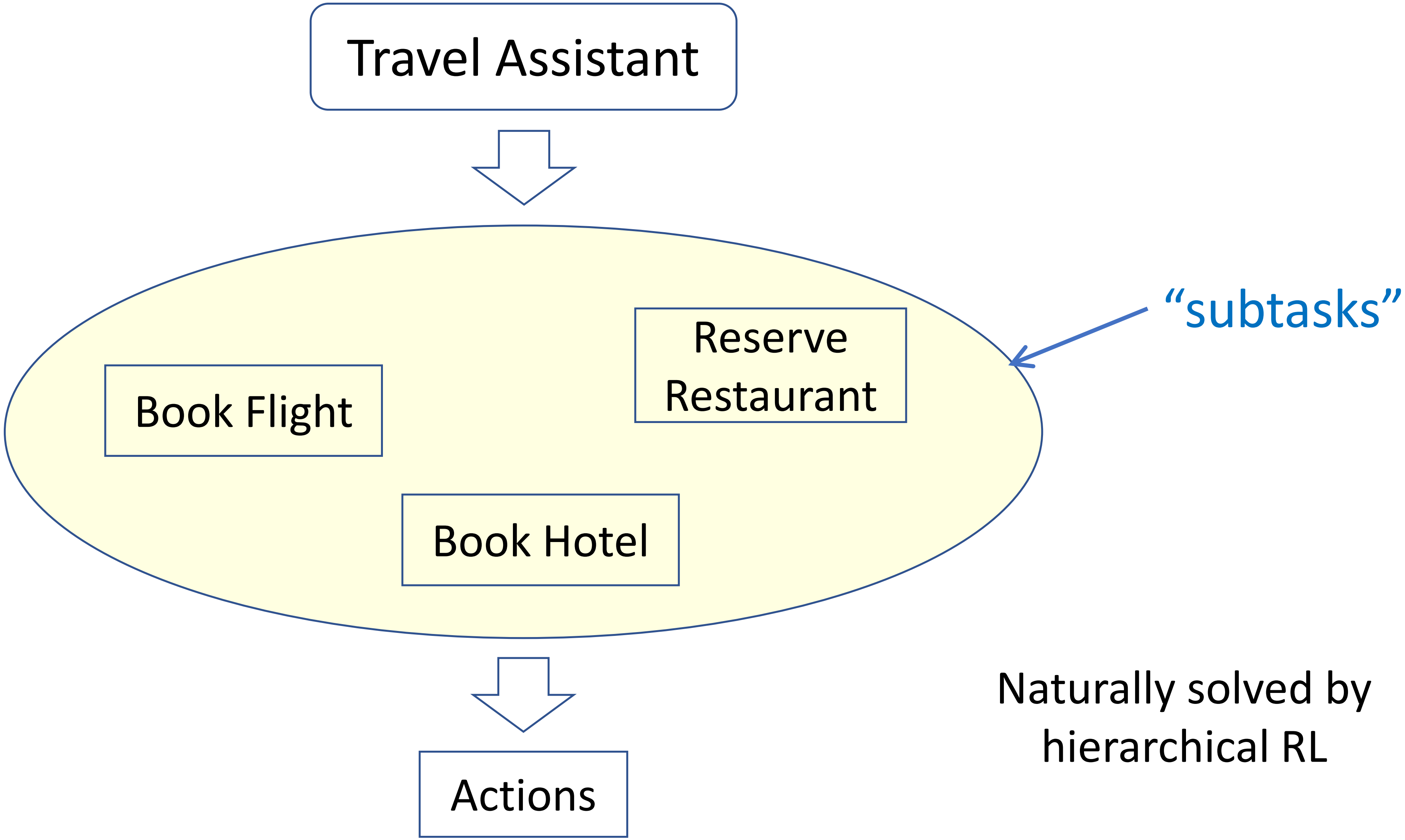



# Three case studies

- Info bots: end-to-end training with non-differentiable knowledge base
- **Composite task completion bots with Hierarchical RL [Peng+ 17]**
- **Task-completion bots: efficient exploration for domain extension**

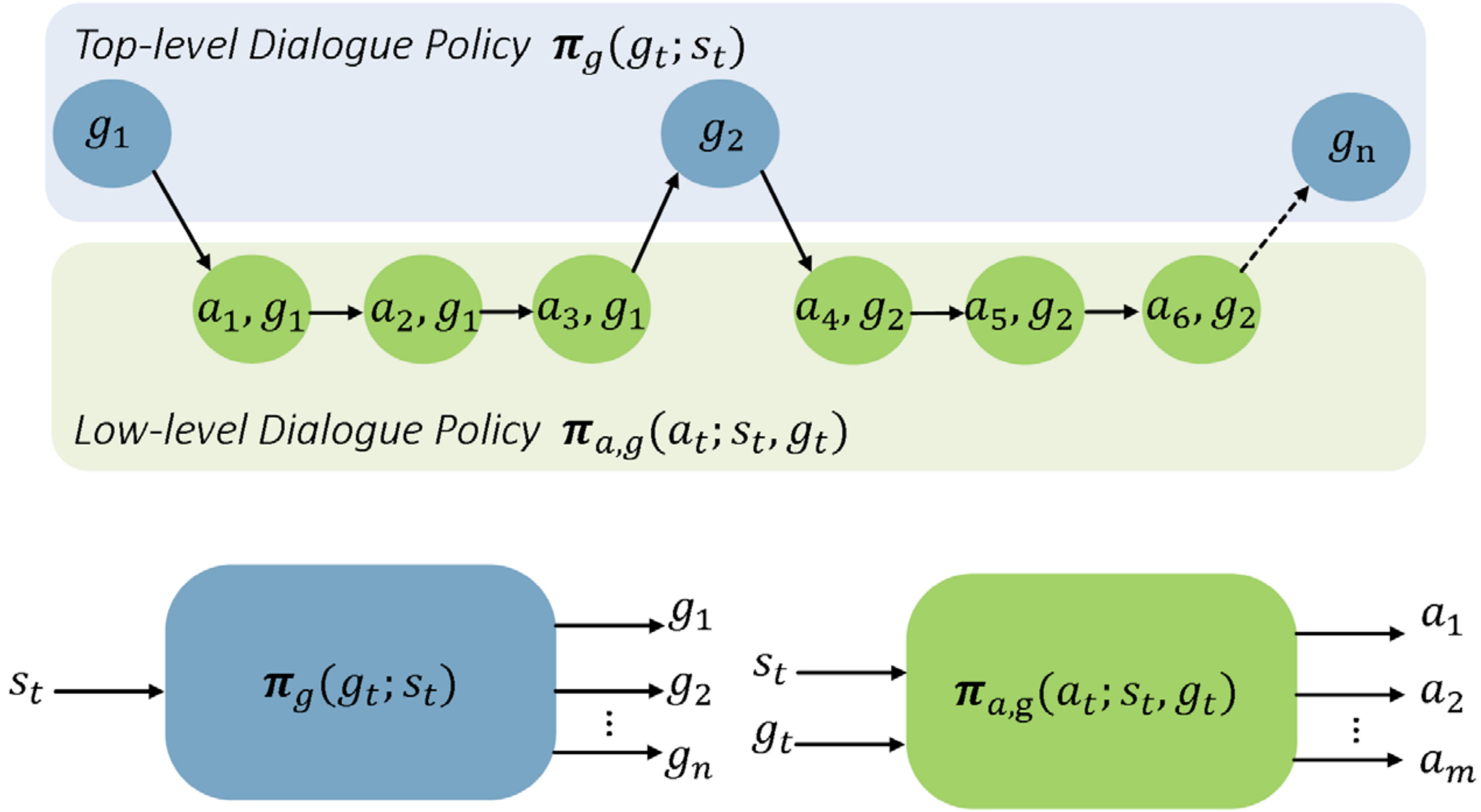
# Composite task completion bot with Hierarchical RL

[Peng+ 17]



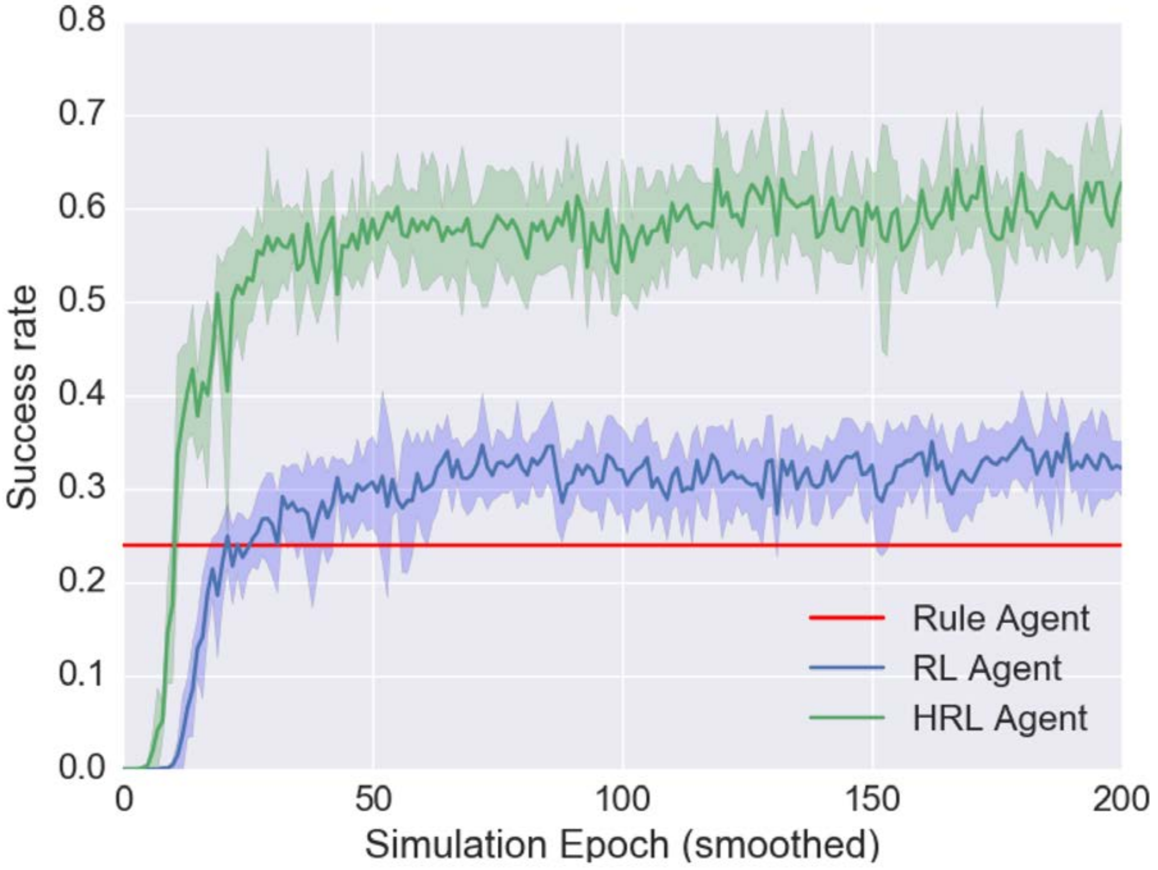

Naturally solved by hierarchical RL

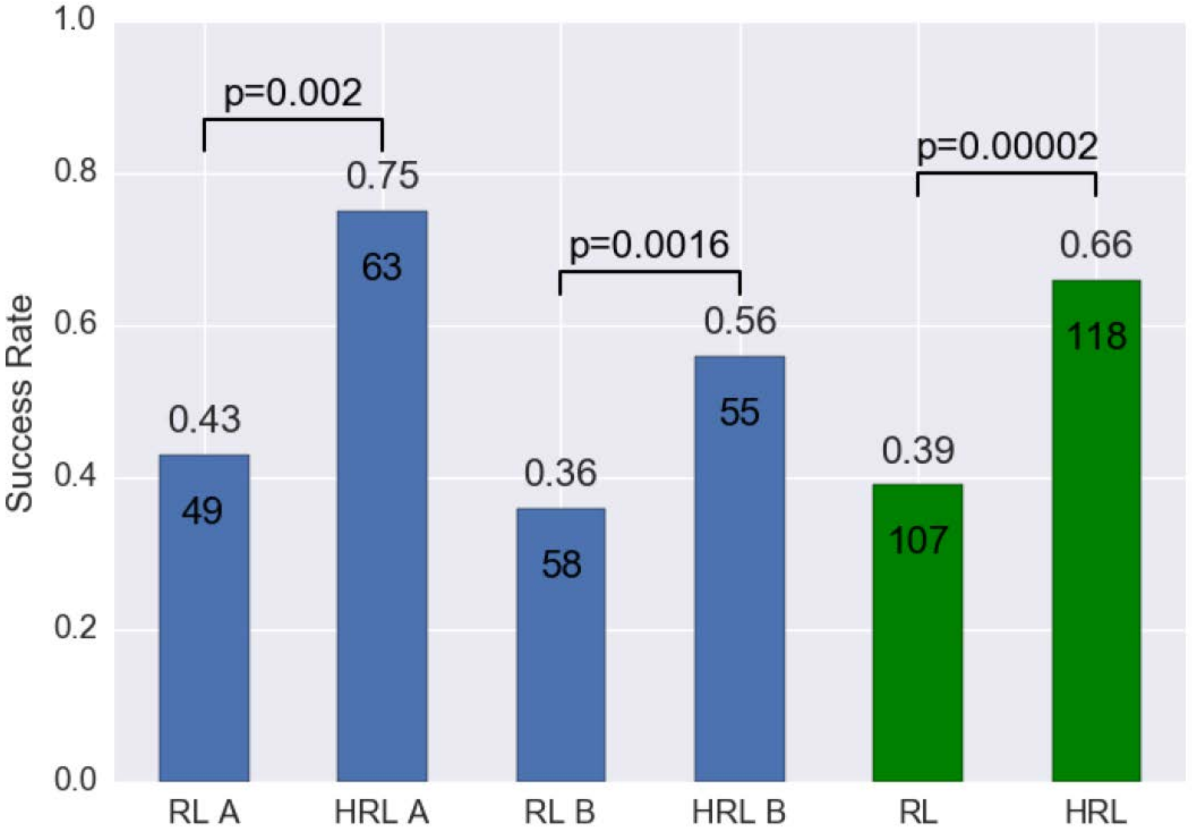
# A hierarchical policy learner




Similar to HAM [Parr & Russell 98] and hierarchical DQN [Kulkarni+ 16]

# Results on simulated and real users



# Subgoal discovery for HRL:

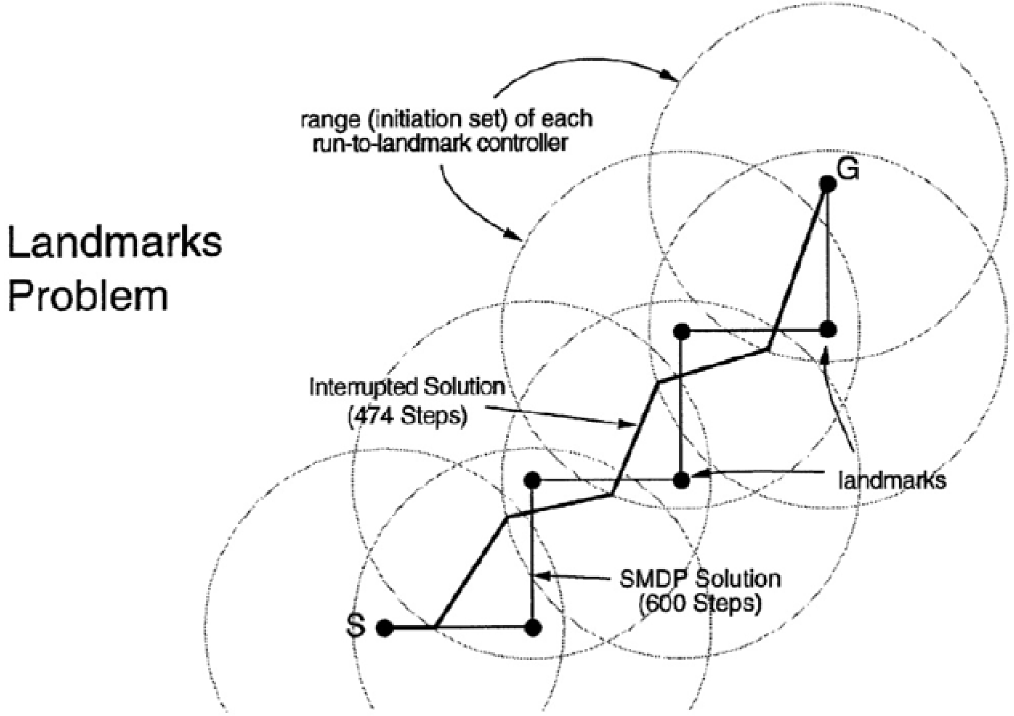



Figure 3: Subgoals for the landmarks problem (Sutton et al., 1999). Though the solution with subgoals may not be optimal, having the subgoals could usually reduce the search space, and potentially accelerate the learning efficiency.

divided and conquer

# The 4-room game

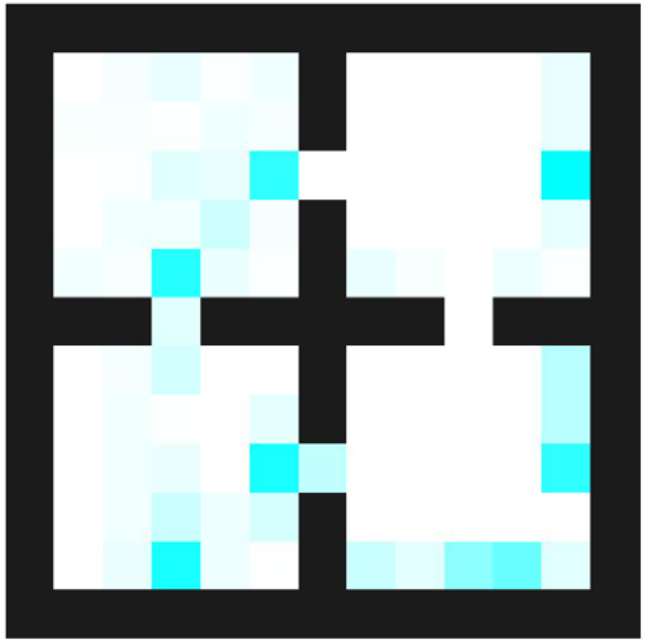


Figure 7: Termination probability visualization for the 4-room experiment. Each time the agent travels from the upper-left corner cell to the lower-right corner cell. The visualization shows the termination probabilities of the RNN generative models in the HRL training after the sequence segmentation process. Darker colors mean higher probabilities.

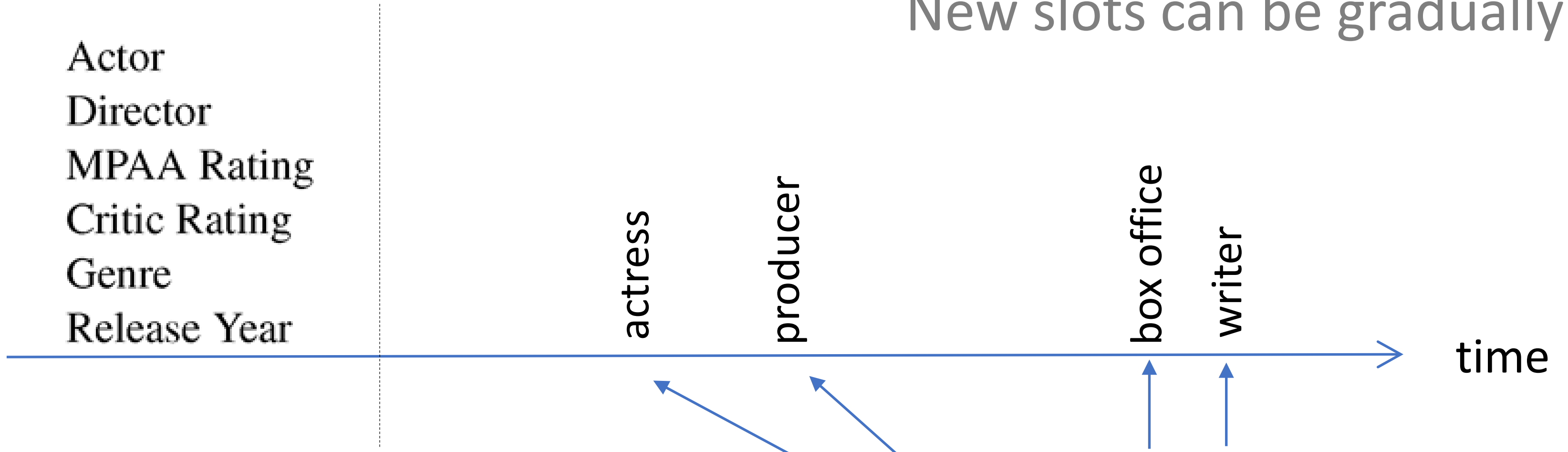
# Three case studies

- Info bots: end-to-end training with non-differentiable knowledge base
- Composite task completion bots with Hierarchical RL [Peng+ 17]
- **Task-completion bots: efficient exploration for domain extension [Zachary+ 17]**

# Domain extension

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

New slots can be gradually introduced



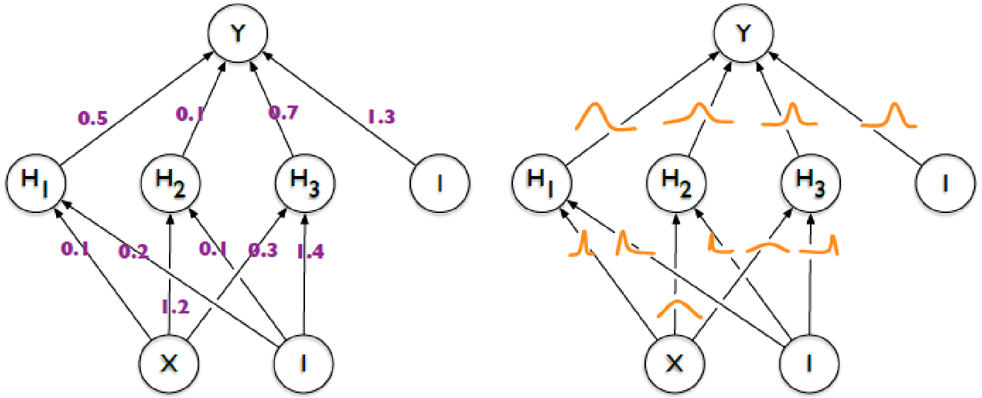

Initial system deployed

**Challenge for exploration:**

- How to explore efficiently
- to collect data for new slots
- When deep models are used

# Efficient exploration for dialogue

- $\epsilon$-greedy can be slow & wasteful, frequently trying known bad moves
  - Compared to Atari/Go settings, failures in dialogue systems confer high economic costs
- Given uncertainty information, one can make smarter exploration decisions
  - DQNs give best estimates of value functions, but don’t offer uncertainty information
- Our solution: get uncertainty info from Bayesian neural networks
  - Explore in area where the model is not confident



# Deep Bayes-by-Backprop Q Network (Deep BBQ Networks)

- Construct a BBQN w. Gaussian variational dist. and Gaussian prior
- Explore by Thompson sampling, drawing Monte Carlo (MC) samples from a stochastic neural net
  - draw $\boldsymbol{w_t}$ from $q(\boldsymbol{w}|\theta)$.
  - set $a_t = \operatorname{argmax}_a Q(s_t, a; \boldsymbol{w}_t)$
- At train time draw one MC sample from BBQN and update by SGVB, using the re-parameterization trick [Kingma & Welling 13]

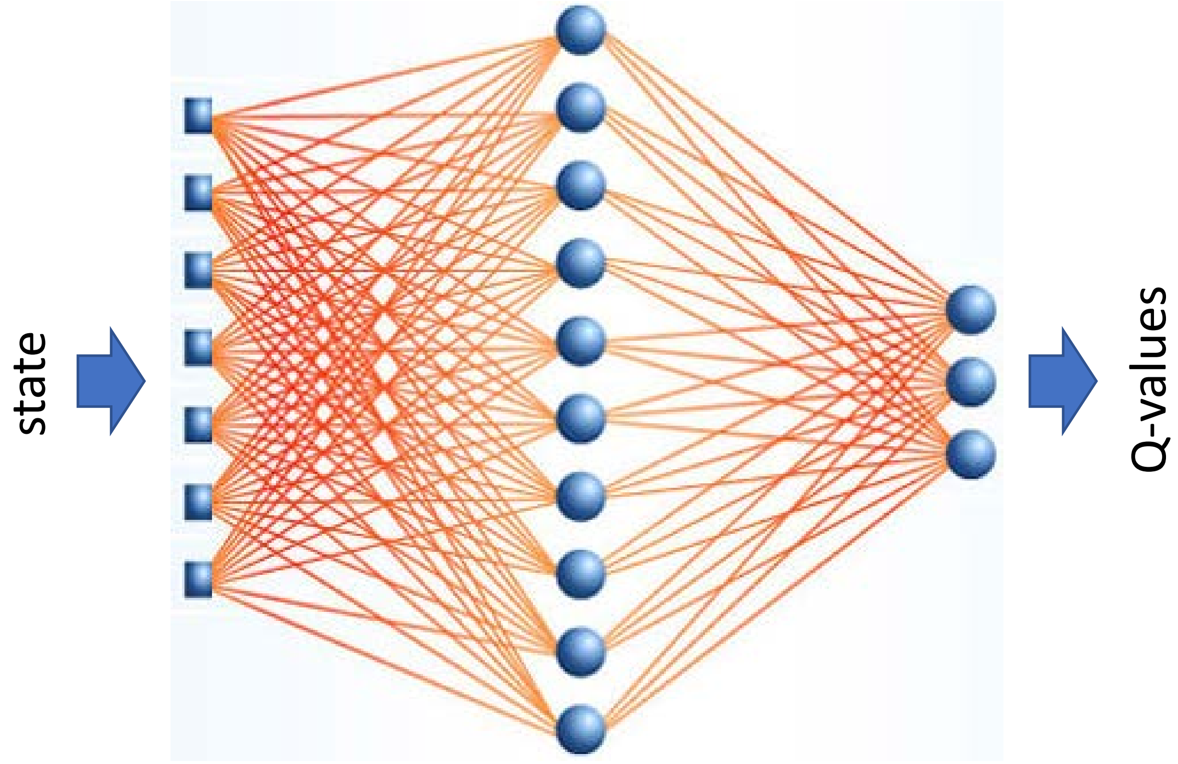
# Deep Q-network (DQN)

DQN-learning of network weights $\theta$: apply SGD to solve

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

“Target network” to
synthesize regression target

“Learning network” whose
weights are to be updated



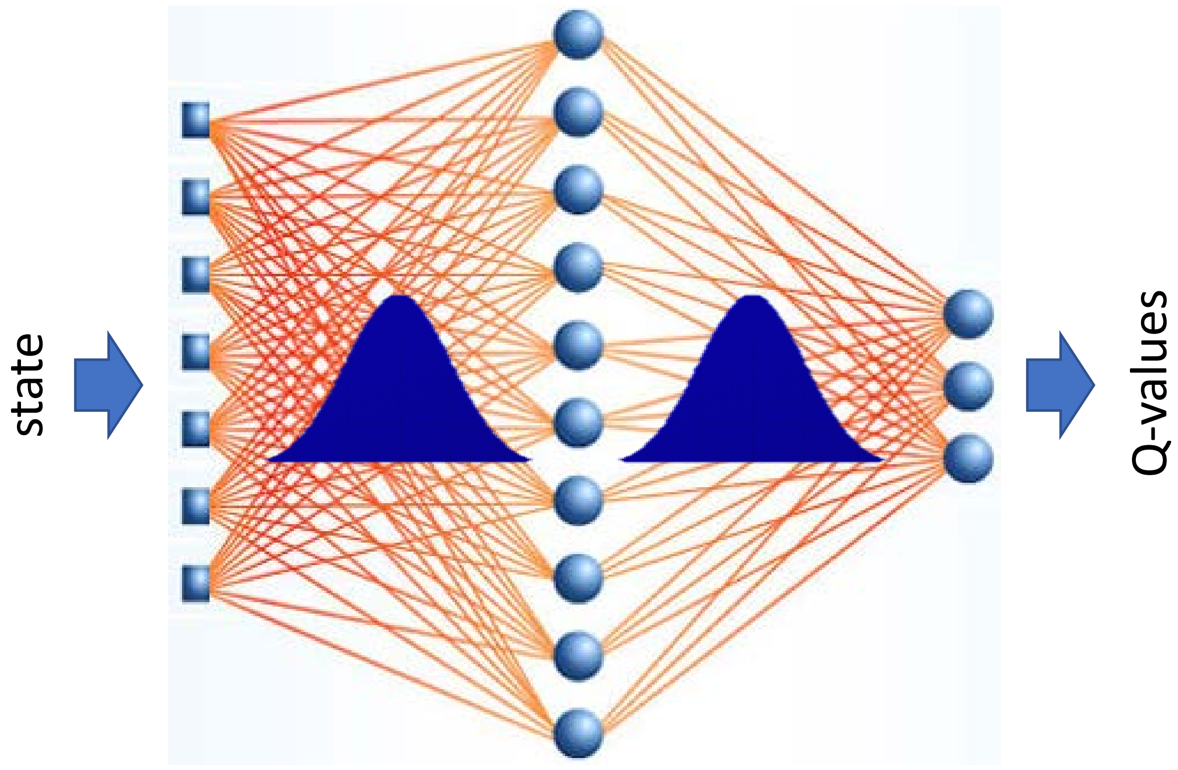

[DeepMind 15]

# Bayes-by-Backprop Q (BBQ) network

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

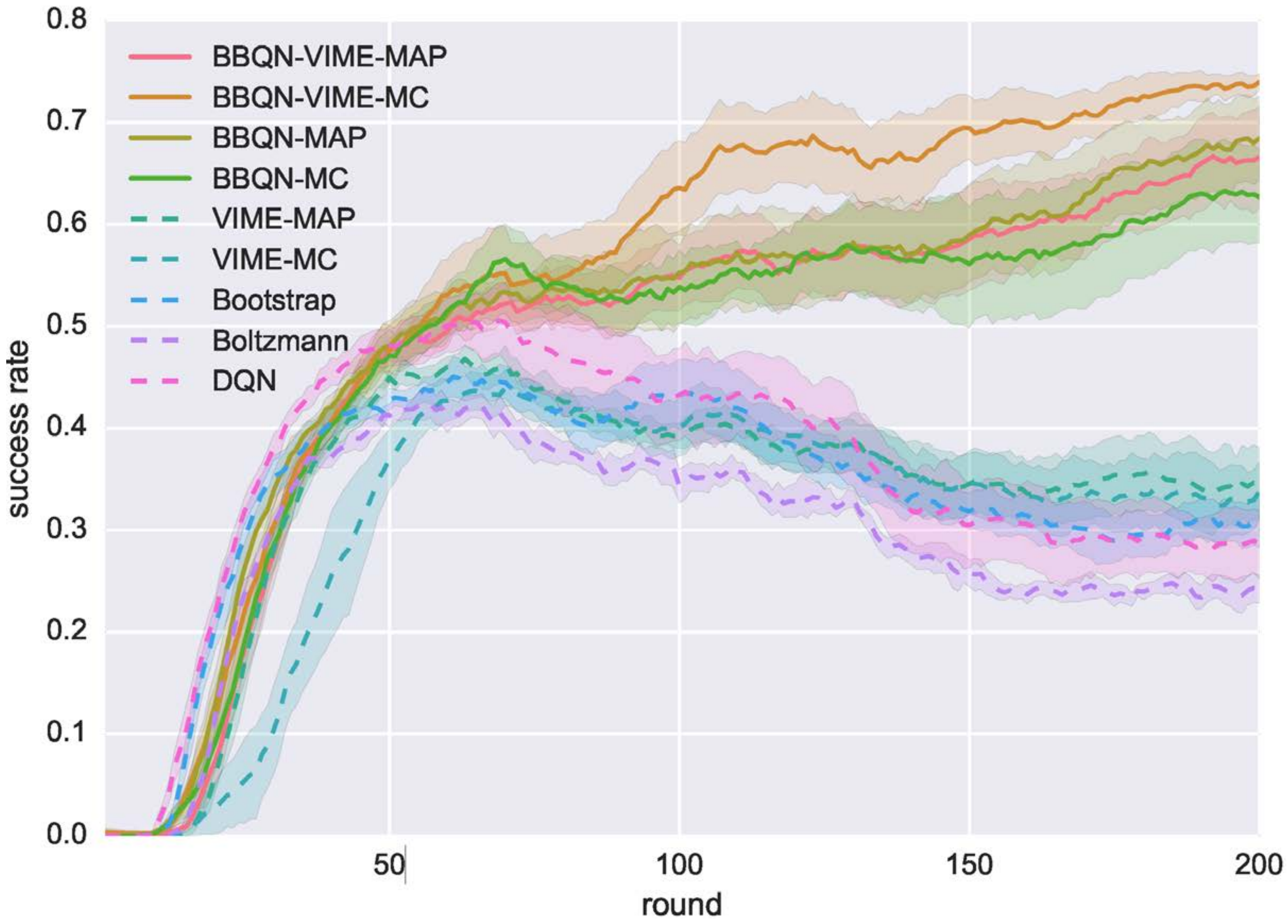
$$\hat{\theta} = \arg\min_{\theta_L} \mathrm{KL}\big(q(\mathbf{w}|\theta_L) \,||\, p(\mathbf{w}|Data)\big)$$

Still use "target network" $\theta_T$
to synthesize regression target




- Parameter learning: solve for $\hat{\theta}$ with Bayes-by-backprop [Blundell+ 15]
- Params $\theta$ quantifies uncertainty in Q-values
- Action selection: use Thompson sampling for exploration

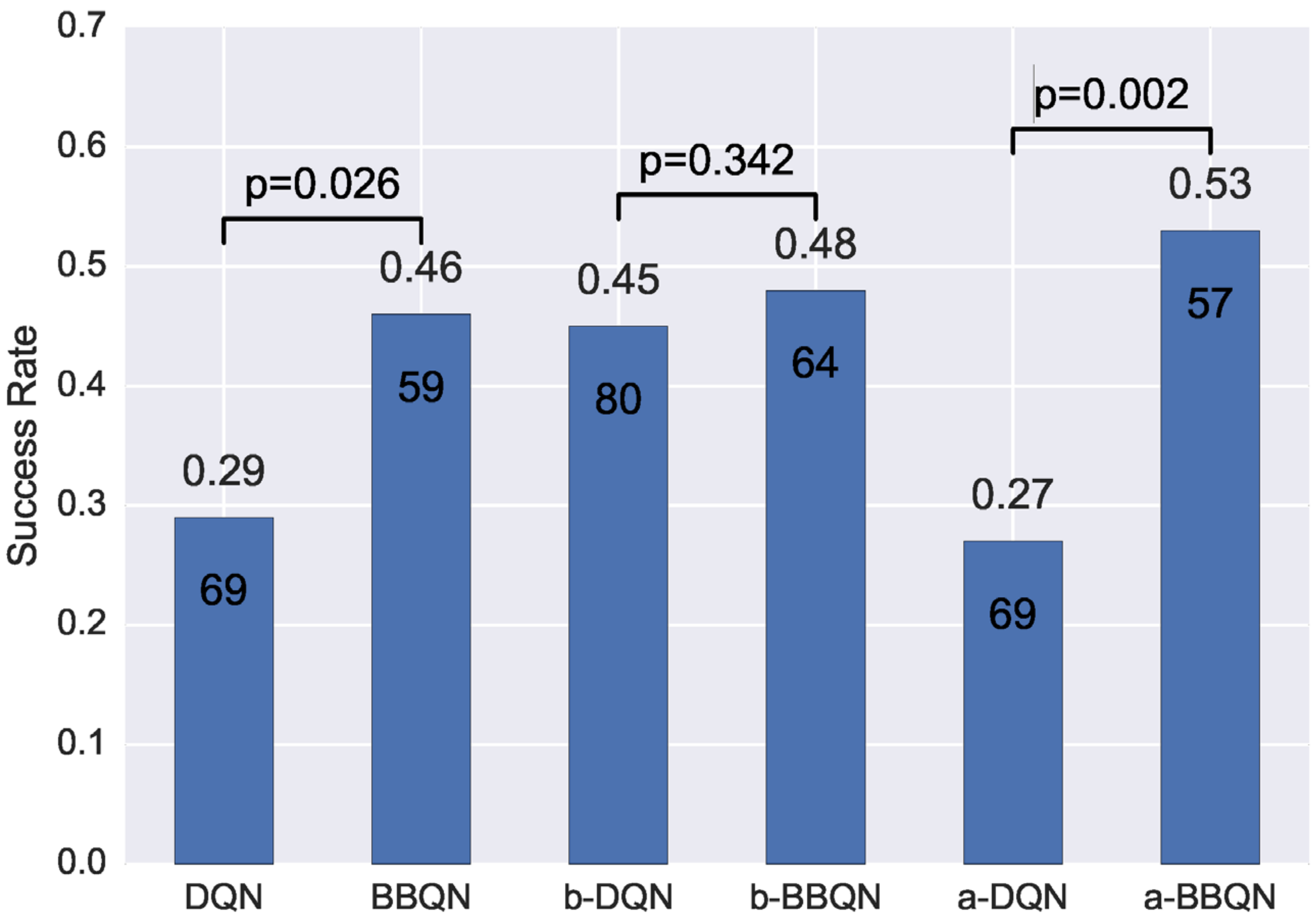
# Results on simulated users




Our BBQ approach successfully explores to adapt to handle new slots.

It also works best in regular dialogue settings (with fixed/full domain)

# BBQ results with real users




- **DQN/BBQN**: regular dialogue policy learning (with full/fixed domain)
- **b-***: model trained on smaller domain
- **a-***: models trained after domain extension

# Summary

- Neural approaches to MRC and QA
  - Knowledge representation and search in neural space
  - A case study: ReasoNet w/ long-term memory
  - Ongoing research: visualize the reasoning process in neural space
  - Learn more at Deep Learning for Machine Reading Comprehension

- An intelligent, human-like, open-domain conversational system
  - Dialogue as RL
  - Two case studies: Info bot and composite-task completion bot
  - Ongoing research: subgoal discovery for hierarchical RL
  - Learn more at deep RL for goal-oriented dialogues