Learning to Understand Questions and Organize Knowledge

Ni Lao

8.25.2021

Speaker Background

Research Scientist @Apple	• Web Answers, KG Answers
Chief Scientist @mosaix.ai	• Created the AI platform team for NLP/ML/quality infrastructures • Research and publications
Research Scientist @Google	• Model based KG construction/cleaning • Semantic parsing for KG QA • Online/offline methods for Web QA
PhD (ML) @Carnegie Mellon U.	• Large scale inference on KG • Relational/structure learning • IR, NLP, QA
MS (CS) BS (EE) @Tsinghua U.	• Automatic system diagnosis based on IR and data mining • Web search, Product search • World champion of RoboCup simulation league (2001 & 2002)

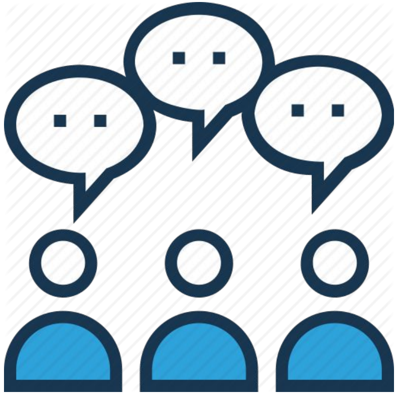
Plan

- ***Query understanding***
  - Weak supervision semantic parsing tasks
  - Neural Symbolic Machines
    - Symbolic representations for efficient inference
  - Memory Augmented Policy Optimization
    - Unbiased low-variance gradient estimation with experience replays
    - RL vs MML vs ML
  - Reranking for RL trained decoders
    - Sequence scorer, Stacked Learning, Scorer Ensembles

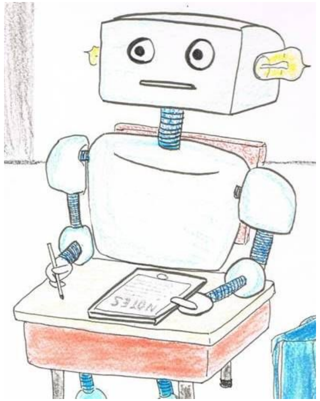
- ***Document understanding***
  - Generalizable, yet accountable & scalable
  - The return of inverted lists
  - Experiment with n-gram machines

Weak Supervision Query Understanding

- Reducing the knowledge acquisition bottleneck with ML is challenging
- Both ML and data labeling/acquisition are central to system design



end to end examples
(QA pairs, preferences)



machine learning

intelligent systems
with knowledge

[Berant et al, 2013; Yih et al, 2016]

WebQuestionsSP

- 5,810 questions from Google Suggest API & Amazon MTurk[1]
- Answerable through FreeBase, a large open knowledge graph (KG)
- 3,098 training examples, 1,639 testing examples remaining
- Open-domain and contains grammatical error
- Multiple entities as answer => macro-averaged F1

• What do Michelle Obama do for a living?	writer, lawyer
• What character did Natalie Portman play in Star Wars?	Padme Amidala
• What currency do you use in Costa Rica?	Costa Rican colon
• What did Obama study in school?	political science
• What killed Sammy Davis Jr?	throat cancer

[Pasupat & Liang, 2015]

WikiTableQuestions

Breadth

- No fixed schema: Tables at test time are not seen during training, needs to generalize based on column name.

Depth

- More compositional questions, thus require longer programs
- More operations like arithmetic operations and aggregation operations

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...	...	...	...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

x_1: *"Greece held its last Summer Olympics in which year?"*
y_1: $\{2004\}$

x_2: *"In which city's the first time with at least 20 nations?"*
y_2: $\{\text{Paris}\}$

x_3: *"Which years have the most participating countries?"*
y_3: $\{2008, 2012\}$

x_4: *"How many events were in Athens, Greece?"*
y_4: $\{2\}$

x_5: *"How many more participants were there in 1900 than in the first year?"*
y_5: $\{10\}$

Domain Specific Languages

[Liang+ 2017]

- An interpreter for LISP like syntax
  - Program => exp_1 exp_2 … exp_n <END>
  - Exp => (f arg_1 arg_2 … arg_n)

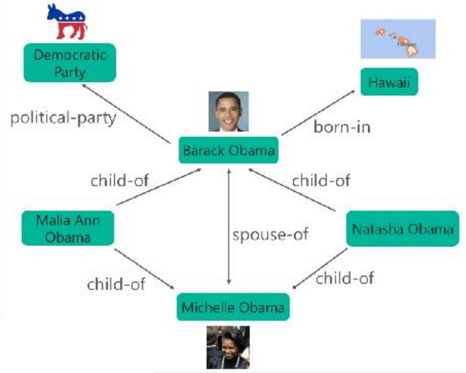
```
(hop m.russell_wilon /education)
(hop v0 /institution)
(filter= v1 m.univeristy
          /notable_types)
<END>
```

- Functions
  - 4 operations for WebQuesitonsSP
    - hop, argmax, argmin, filter
  - 22 different operations for WikiTableQuestions
    - hop, argmax, argmin
    - $filter_{=}$, $filter_{!=}$, $filter_{>}$, $filter_{<}$, $filter_{>=}$, $filter_{<=}$, $filter_{in}$, $filter_{!in}$,
    - first, last, previous, next
    - max, min, average, sum, mode, diff, same

Semantics as A Foreign Language

Largest city in US?

```
GO
(Hop V1 CityIn)
(Argmax V2 Population)
RETURN
```



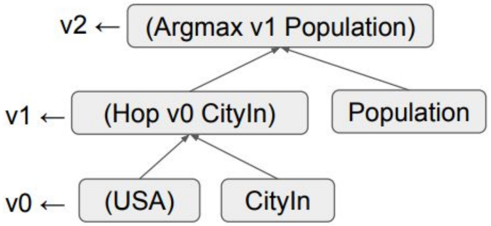
Freebase, DBpedia, YAGO, NELL

NYC

Paraphrase

Many ways to ask the same question, e.g.,
"What was the date that Minnesota became a state?"
"When was the state Minnesota created?"

Compositionality



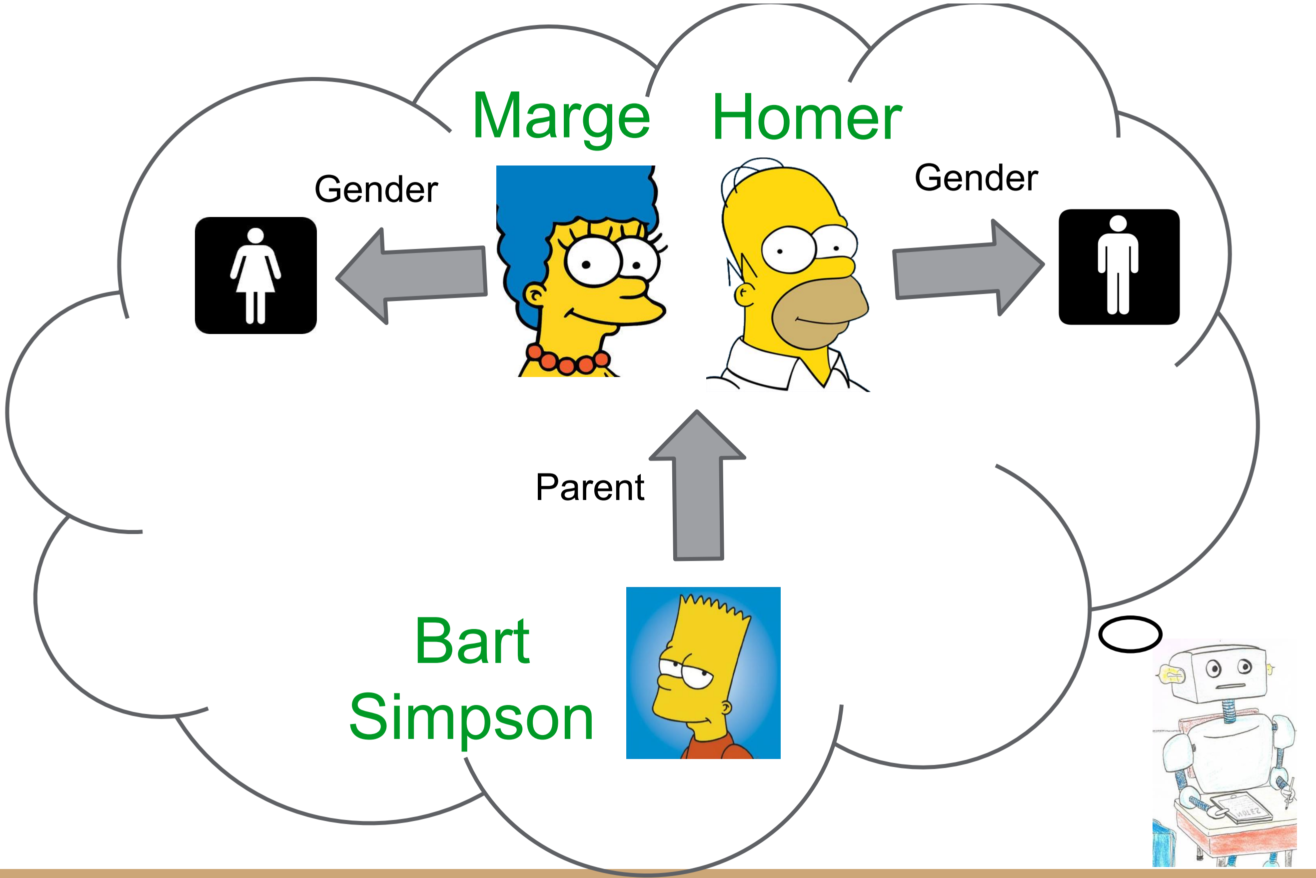
Large Search Space

E.g., Freebase:
23K predicates,
82M entities,
417M triplets

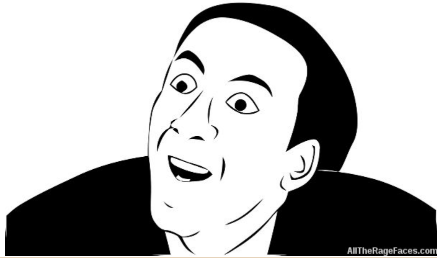
Optimization

Directly optimizing the metric (e.g., F1) with RL faces challenges

Understanding Needs Reasoning



Bart's father is Homer



Reasoning May Incur A Lot of Computation

“impressionist painters during the 1920s”

painters	[/painting] !/art_forms
impressionist <visual_artist>	x.[/associated_periods_or_movements = /impressionism]
<artist> during the 1920s	x.[/date_of_work < 1930; /date_of_work > 1920]

Symbolic Computations & Neural Nets

- Combine the learnability of neural nets with the efficiency of symbolic reasoning

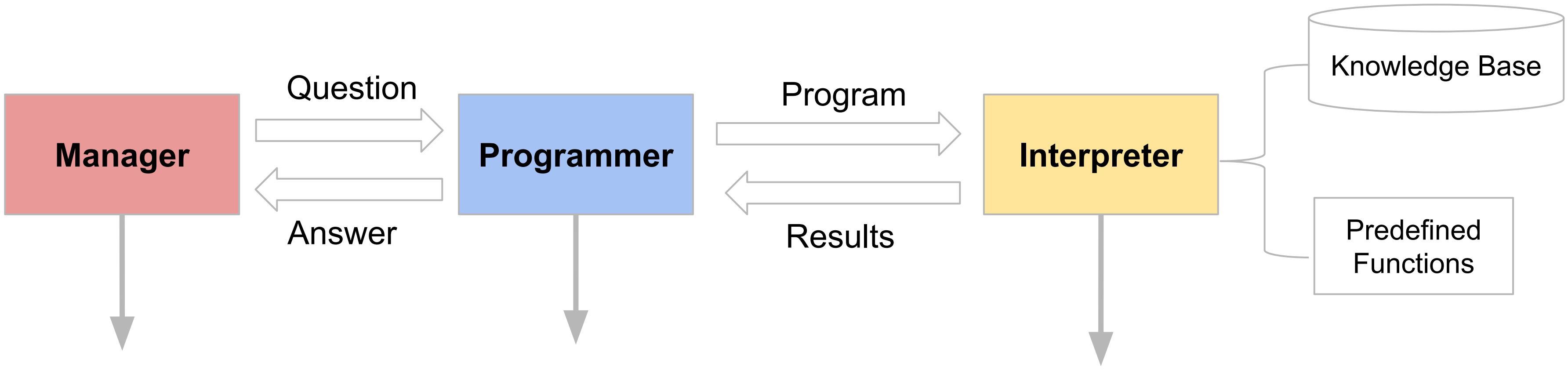
a symbolic machine

a neural controller

a sequence of symbols

Neural Symbolic Machines

[Liang+ 2017]



Weak supervision

Low cost
User centric

Neural Net

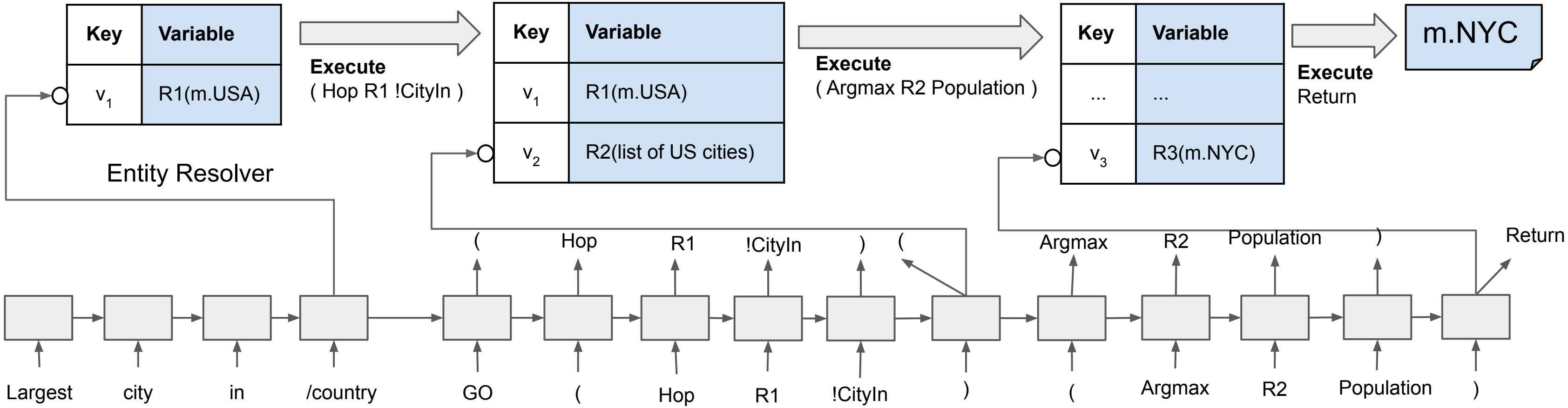
Learning power

Symbolic

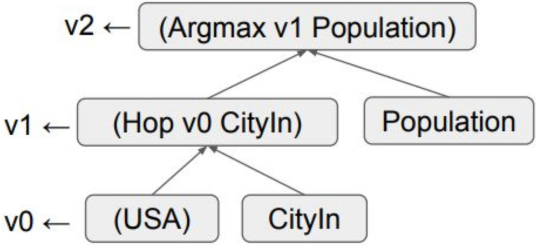
Abstract
Scalable
Precise
Non-differentiable

Seq2Seq with Variables for Compositionality

[Liang+ 2017]



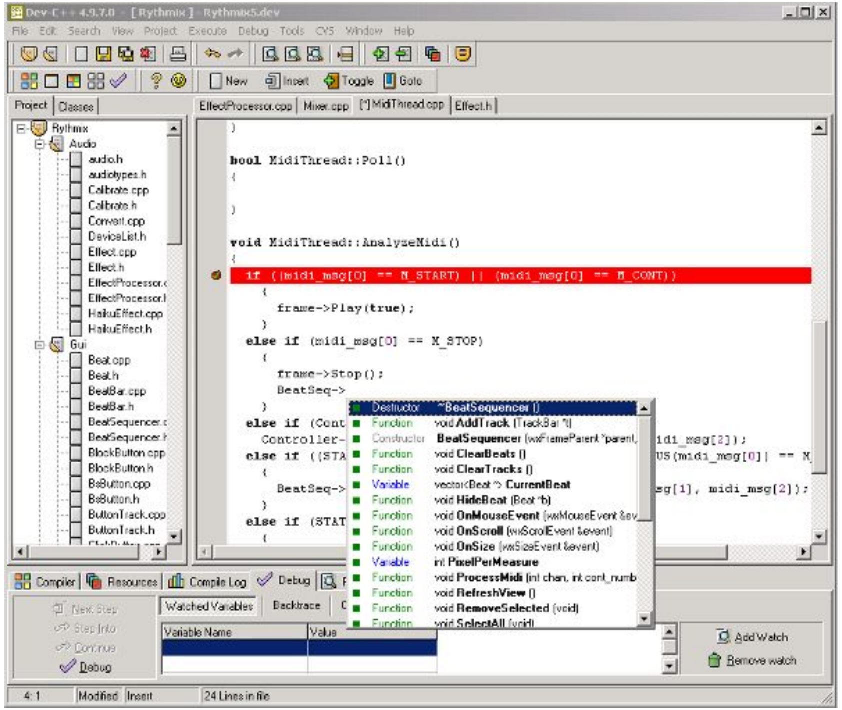
- Equivalent to a linearised bottom-up derivation of the recursive program
- Aggressive pruning by code assists



[Liang+ 2017]

Code Assistance with An Interpreter

Pen and paper

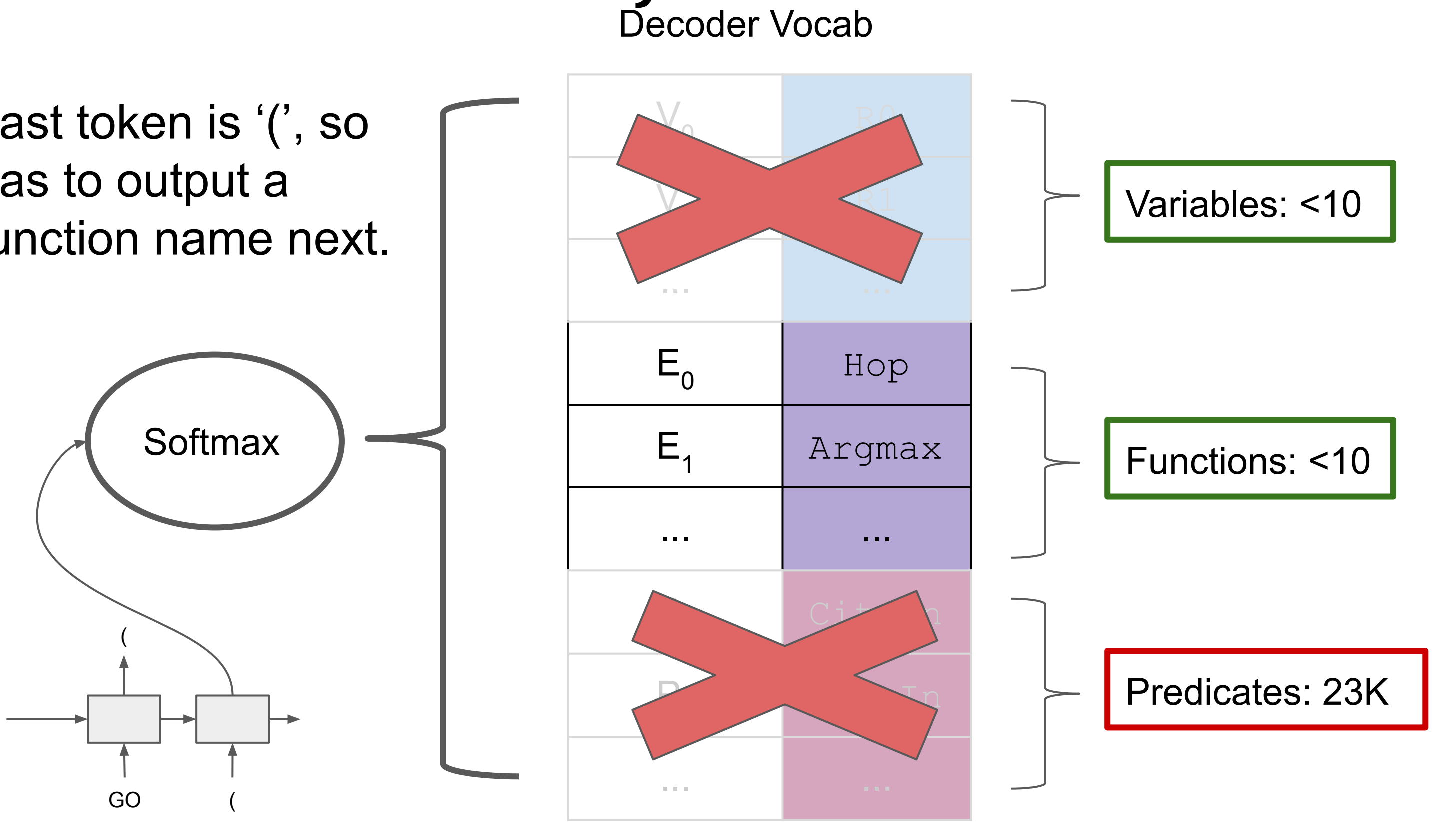


IDEs do a lot of computations!

[Liang+ 2017]

Code Assistance: Syntactic Constraint

Last token is ‘(’, so has to output a function name next.

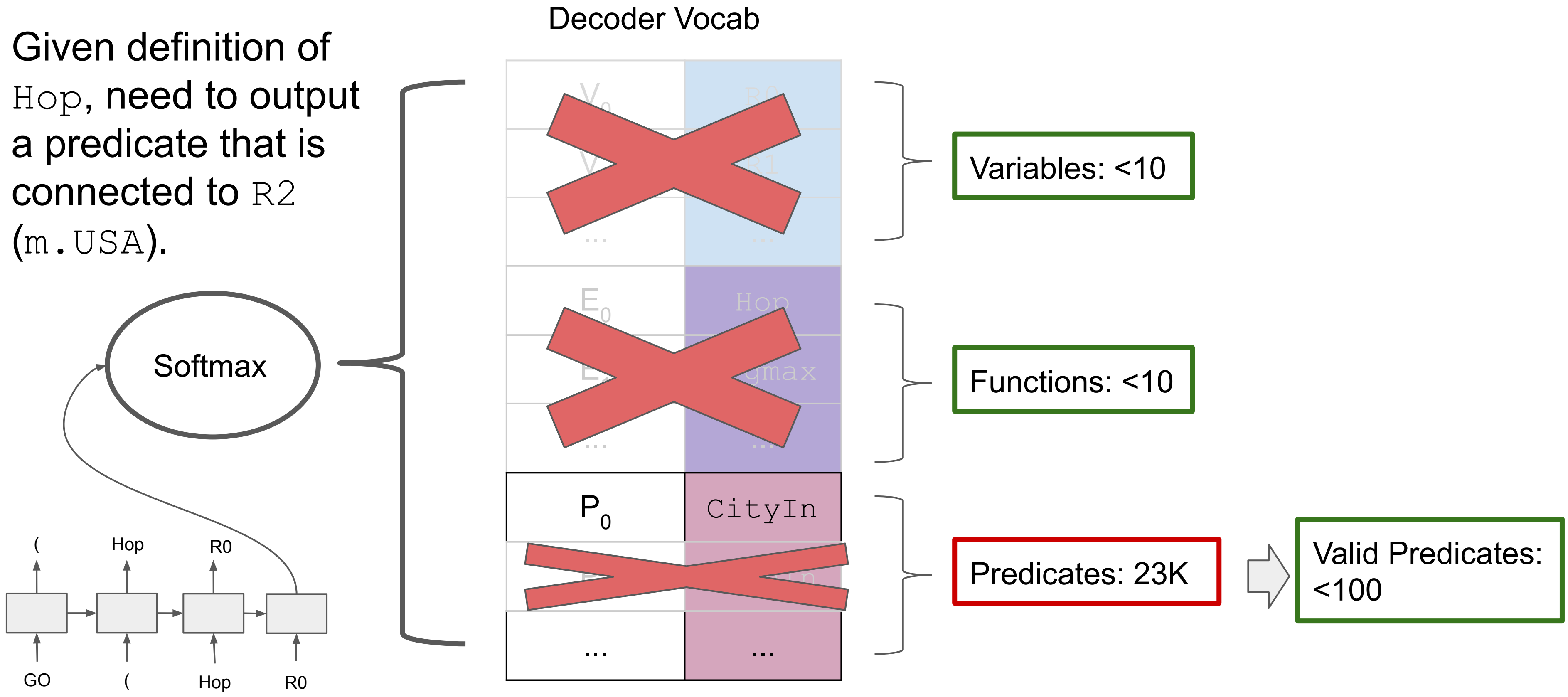


Details

Code Assistance: Semantic Constraint

[Liang+ 2017]

Given definition of `Hop`, need to output a predicate that is connected to `R2` (`m.USA`).



Directly Optimizing The Expected Reward with RL

[Sutton & Barto 1998, 2018]

- ML optimizes the log likelihood of target sequences

$$J^{ML}(\theta) = \sum_q \log P(a_{0:T}^{best}(q) | q, \theta)$$

- RL optimizes the expected reward under a stochastic policy

$$J^{RL}(\theta) = \sum_q \mathbb{E}_{P(a_{0:T}|q,\theta)}[R(q, a_{0:T})]$$

Augmented REINFORCE

[Liang+ 2017]

- Iterative ML training is not directly optimizing the F1 scores
- REINFORCE get stuck at local maxima
- Augmented REINFORCE obtains better performances
    - but the objective is biased

WebQuestionsSP Results

Settings	Train Avg. F1@1	Valid Avg. F1@1
iterative ML only	68.6	60.1
REINFORCE only	55.1	47.8
Augmented REINFORCE	83.0	**67.2**

[Liang+ 2016]
[Abolafia+ 2017]

Augment REINFORCE

Linear combination of ML and RL objective:

- Converges fast to a reasonable policy
- The gradient is biased -- not robust against spurious programs

$$\lambda \sum_{y \in \mathrm{TopK}} \log p(y \mid x) + (1-\lambda)\mathbb{E}_{\tilde{y} \sim p(y|x)} R(\tilde{y})$$

highest rewarded solutions in memory

new samples from the current policy

Details

[Liang+ 2018]

Spurious Rewards

Which nation won the most silver medal?

Rank	Nation	Gold	Silver	Bronze	Total
1	Nigeria	14	12	9	35
2	Algeria	9	4	4	17
3	Kenya	8	11	4	23
4	Ethiopia	2	4	7	13
5	Ghana	2	2	2	6
6	Ivory Coast	2	1	3	6
7	Egypt	2	1	0	3
8	Senegal	1	1	5	7
9	Morocco	1	1	1	3
10	Tunisia	0	3	1	4
11	Madagascar	0	1	1	2
12	Rwanda	0	0	1	1
12	Zimbabwe	0	0	1	1
12	Seychelles	0	0	1	1

We can get the right answer with many wrong reasons

- **Correct program**:
  (argmax rows “Silver”)
  (hop v1 “Nation”)

- **Many spurious programs**:
  (argmax rows “Gold”)
  (hop v1 “Nation”)

  (argmax rows “Bronze”)
  (hop v1 “Nation”)

  (argmin rows “Rank”)
  (hop v1 “Nation”)

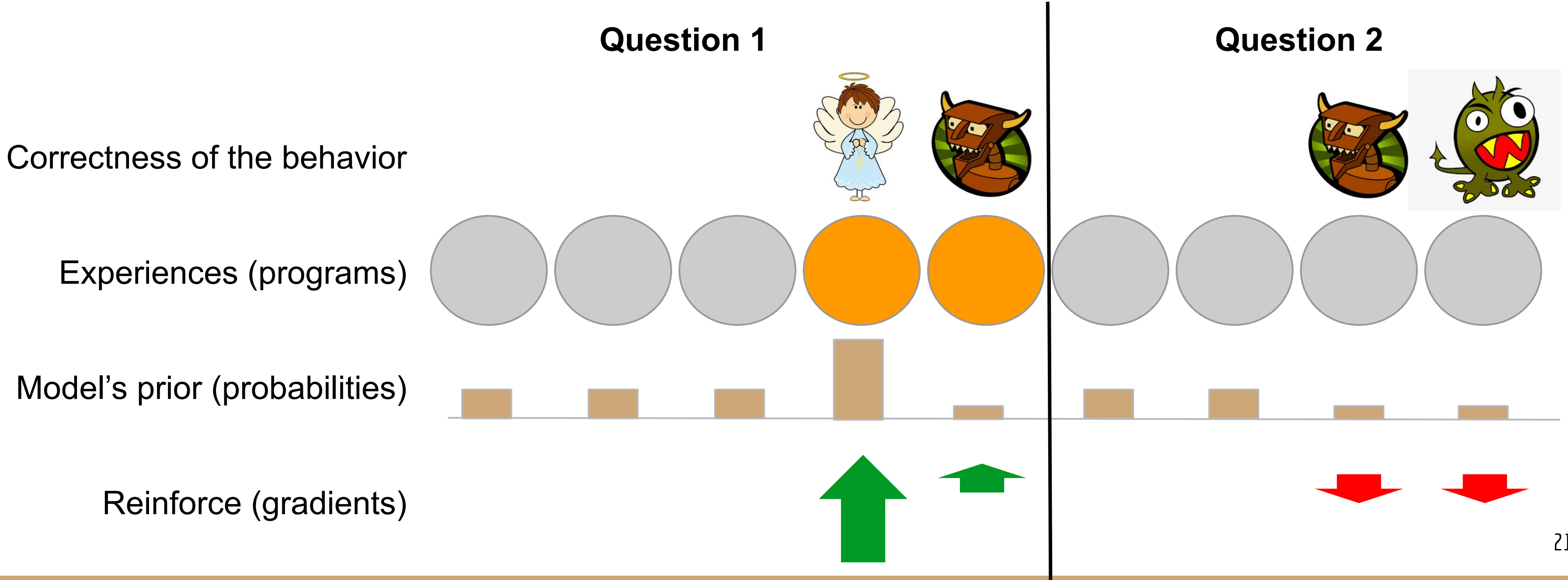
  ...

Details

[Liang+ 2018]

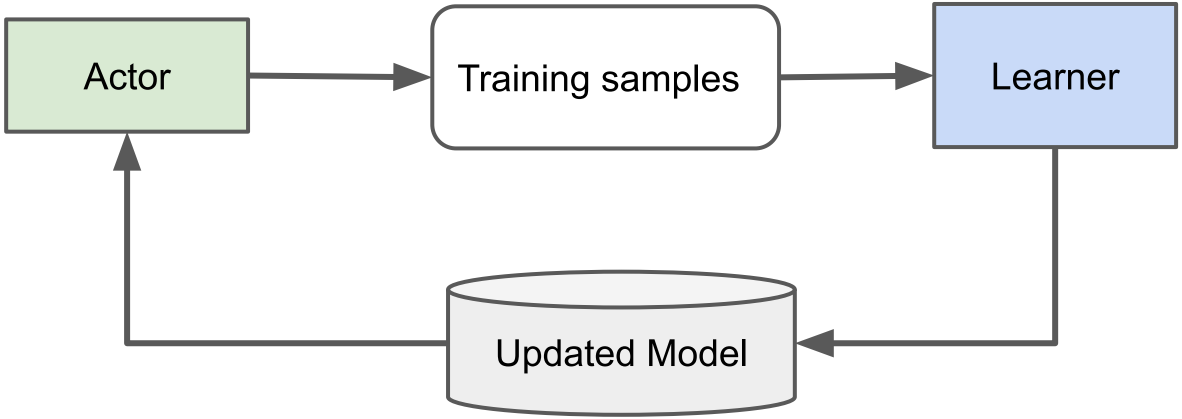
The RL objective demotes spurious rewards

- **Reinforce** a **rewarded experience** only if the model (current policy) also thinks that it is the right thing to do



[Liang+ 2018]

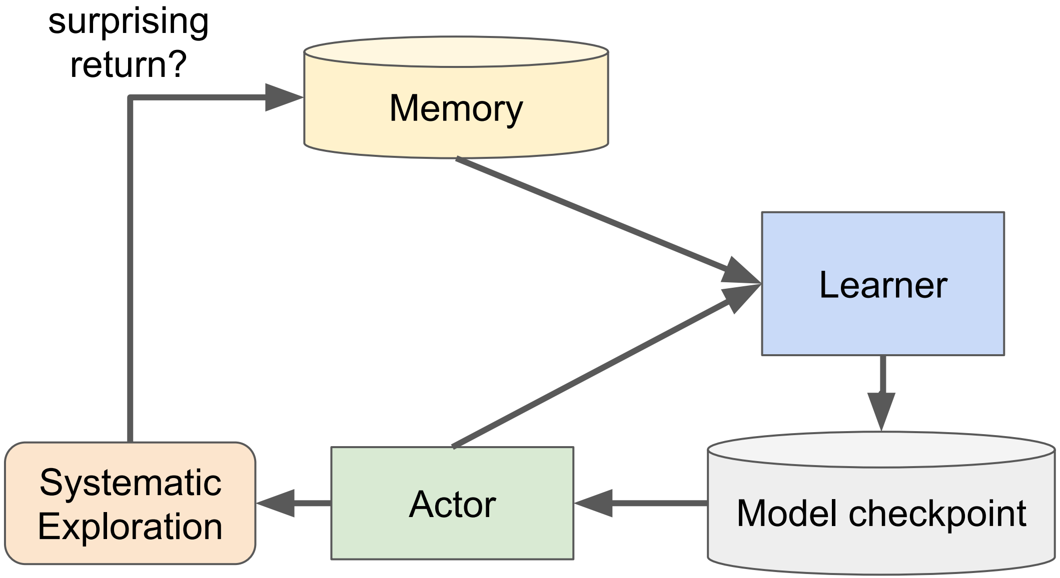
RL models generate their training data on the fly



- Training sample management issue
  - Large search space & sparse reward lead to slow and unstable training
  - Spurious reward lead to biased solutions

[Liang+ 2018]

Memory Augmented Policy Optimization (MAPO)



Most of the past experience are not helpful for improving the current model

http://insideout.wikia.com/wiki/Long_Term_Memory

[Liang+ 2018]

Optimal Sample Allocation

- The optimal strategy (low variance in gradient estimations) is to allocate the **same number of samples** to **reward** vs **no reward** experiences (0-1 reward)
- and this is independent of the model's **current performance**

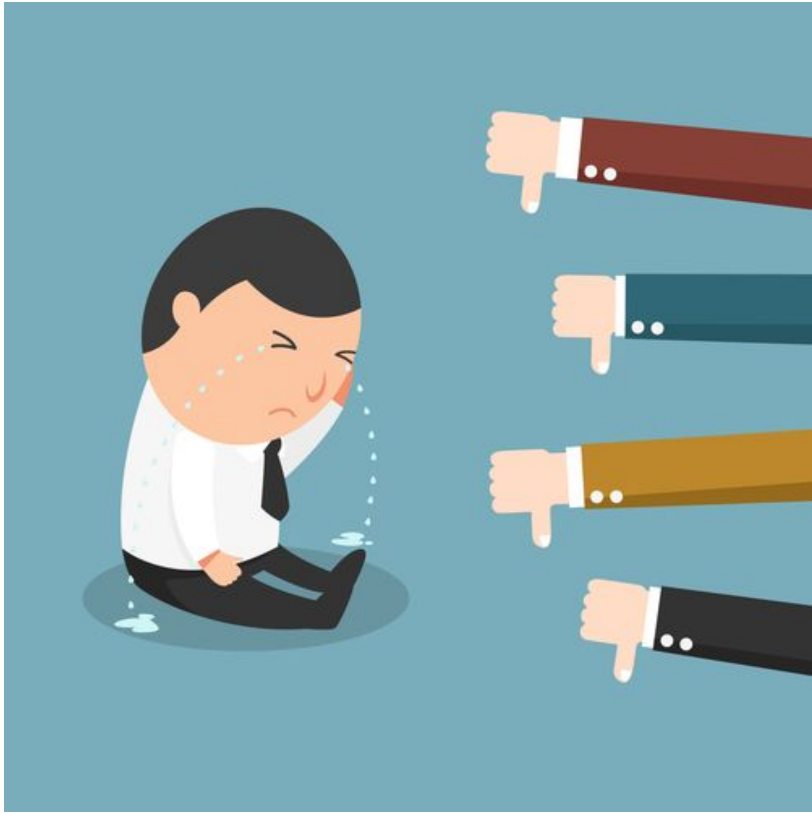


Image source: Guy Harris, 2018
How to Give Feedback in a Non-Threatening Way

[Liang+ 2018]

Unbiased gradient estimation \w low variances

Given a memory buffer of high return sequences $\mathcal{B} \equiv \left\{ \left(y^{(i)}, r^{(i)} \right) \right\}_{i=1}^{n}$,

re-express expected return as,

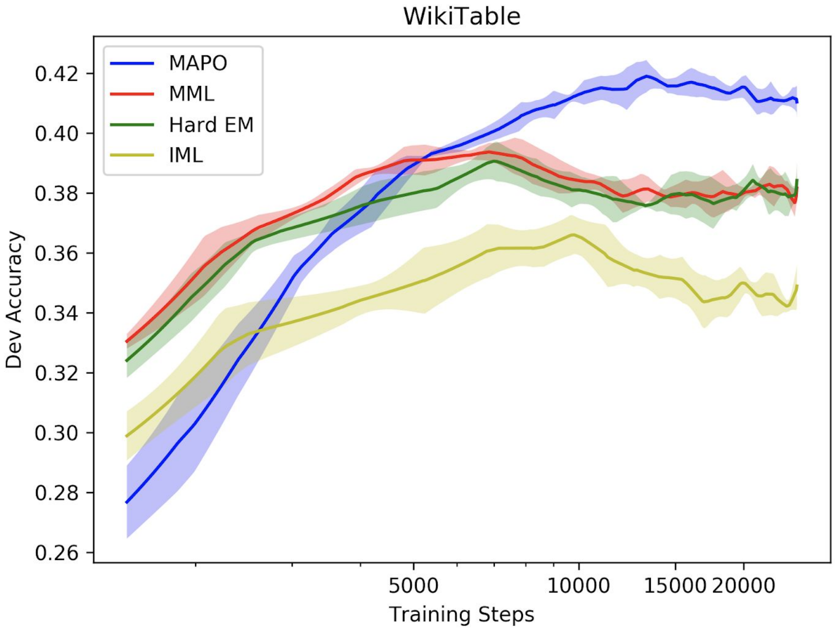
$$p(\mathcal{B}) \underbrace{\mathbb{E}_{p(\tilde{y})|\tilde{y}\in\mathcal{B}} R(\tilde{y})}_{\text{inside the buffer}} + (1 - p(\mathcal{B})) \underbrace{\mathbb{E}_{p(\tilde{y})|\tilde{y}\notin\mathcal{B}} R(\tilde{y})}_{\text{outside the buffer}}$$

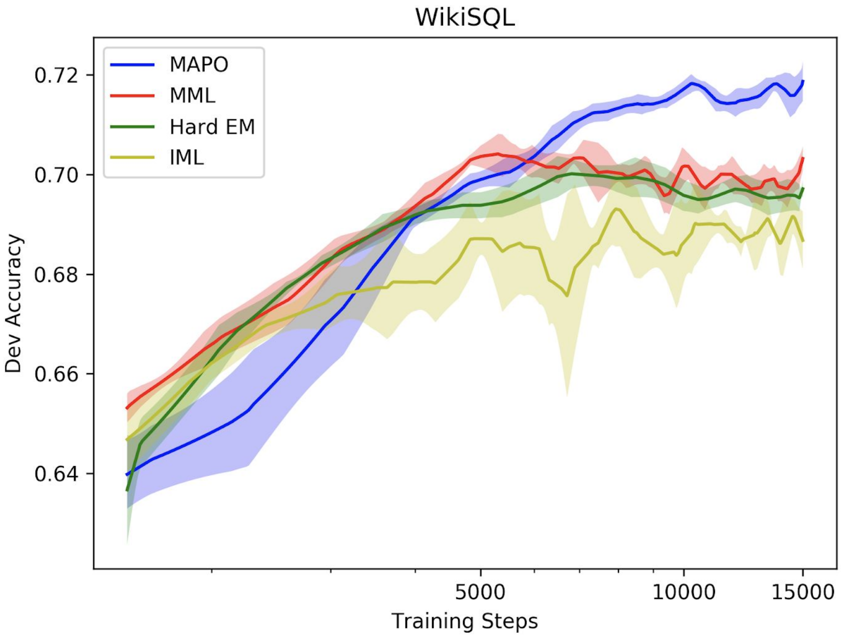
- For each query
  - Sampling 1 solution from inside the buffer according to model
  - Rejection sampling 1 solution from outside the buffer according to model

[Liang+ 2018]

Comparison

- REINFORCE does not work at all
- MAPO is slower but less biased than max marginal likelihood and hard EM



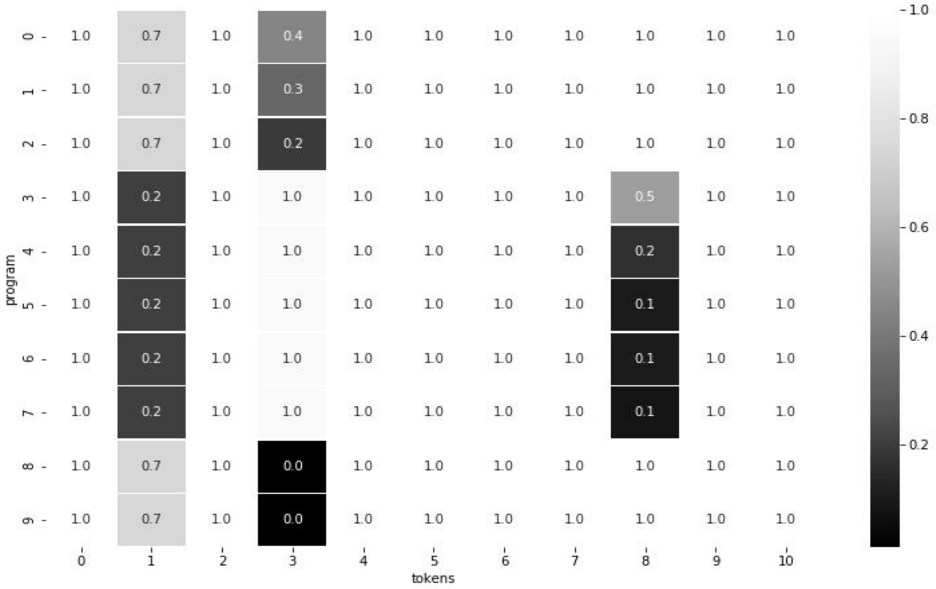


- The shaded area represents the standard deviation of the dev accuracy

The Issues with Sequence Probabilities

[Collins 2000]
[Lafferty+ 2001]
[Biloki+ 2019]

- Cannot consider information in the future; or global statistics
- The label bias problem of sequential models
  - states with limited choices effectively ignore their observations

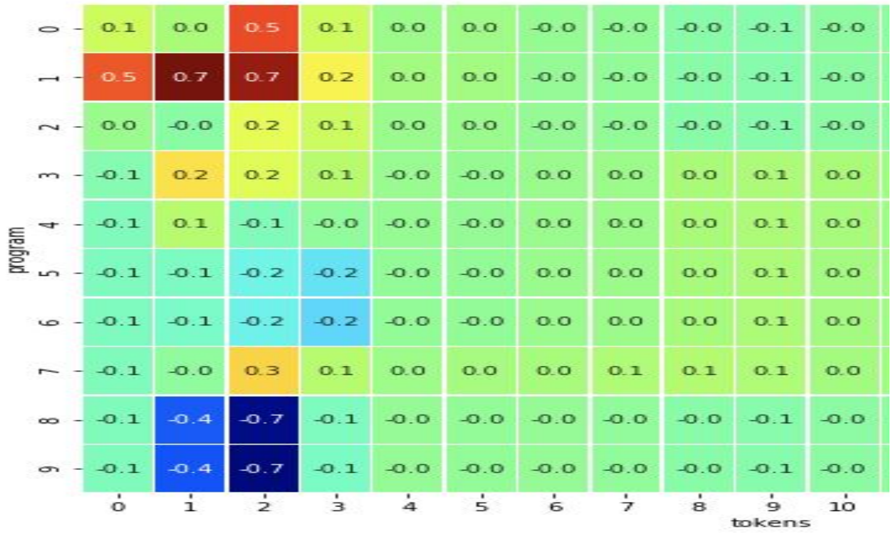


MAPO probability per token for programs in beam. The score sequences have the same length (10) because of padding

Leverage Global Discriminative features

[Biloki+ 2019]

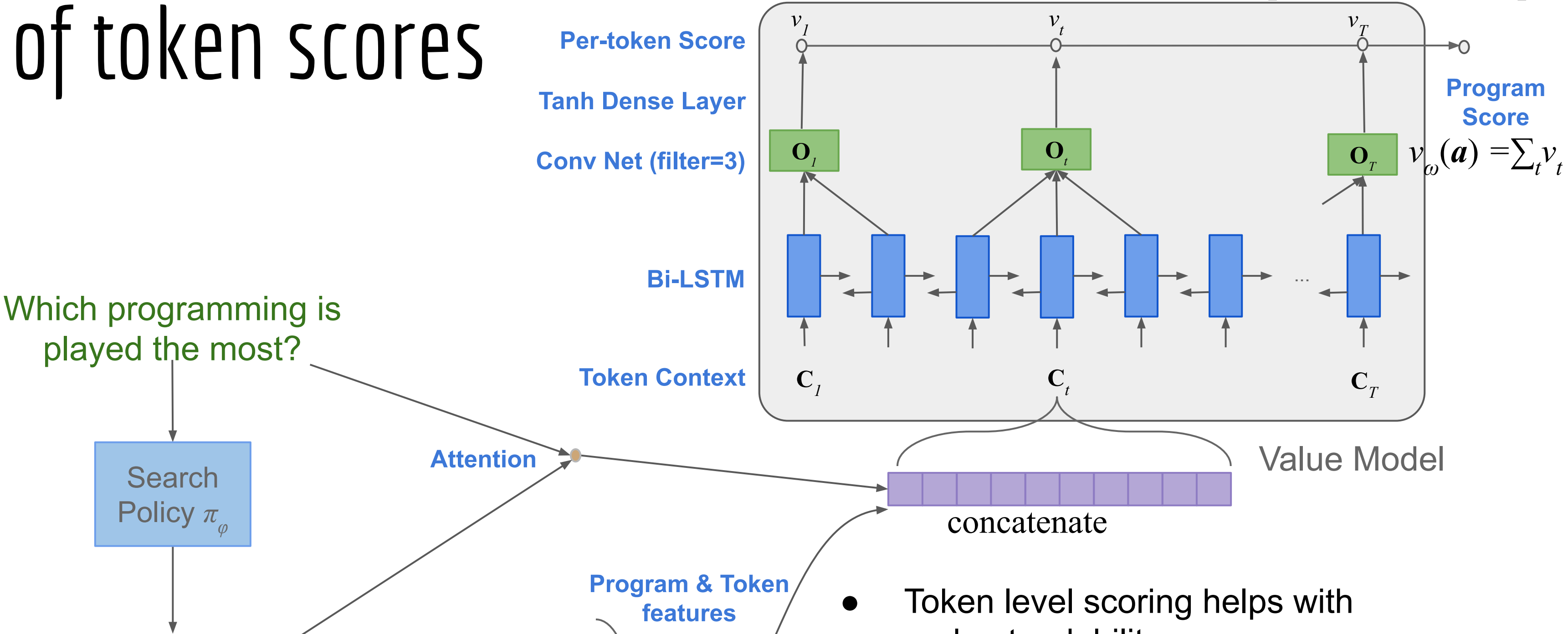
Symbol	Type	Meaning
q^{tok}	binary	The program token matches any of the question tokens.
$\mathbf{q}^{\text{attn}}$	float vector	Softmax attention over query tokens per program token
p^{prob}	float	Program probability according to the search policy π_ϕ
t^{prob}	float	Program token probability according to the search policy π_ϕ
t^{agree}	count	Number of candidate programs having token a_t at position t



NPP scores per token for a set of candidate programs in beam

Sum of token scores

[Biloki+ 2019]



Beam:

(mode all_rows r.location-string) <END> candidate program ***a***

(mode all_rows r.programming-string) <END>

(mode all_rows r.psip-string) <END>

(argmax all_rows r.rf-number) (hop v8 r.location-string)<END>

...

- Token level scoring helps with understandability
- Bi-LSTM considers information forward and backward in time
- ConvNet considers spans equal to the size of lisp clauses

[Biloki+ 2019]

Reranking & Stacked Learning

- **Reranking**
  - train a reranking model (B) to improve the output of a generator model (A)
  - Syntax parsing [Collins 2000; Charniak&Johnson 2005; Huang 2008]
  - Segmentation, POS tagging [Sun 2012]
  - Entity linking [He+ 2013]

- **Stacked learning** to correct the training/test mismatch
  - Create a k-fold cross-validation split on the original training data
  - Train k copies of model (A)
  - Generate the training data for model (B) with these k models
  - Train model (B)

Discriminative Training Objective

- Rank a set of candidates in beam

$$\mathcal{O}_{\text{NPP}}(\omega) = \sum_{l} \sum_{1 \leqslant i \neq j \leqslant |s_\phi(\mathbf{x}^l)|} \mathbb{1}[r^{l,i} > r^{l,j}] \log \sigma(v^{l,i} - v^{l,j})$$

$$\boxed{\sigma(v) = 1/(1 + e^{-v})}$$

- The score of program $\mathbf{a}^{l,i}$ considers query $\mathbf{x}^l$ and beam $s_\phi(\mathbf{x}^l)$

$$v^{l,i} = v_\omega(\mathbf{a}^{l,i}; \mathbf{x}^l, s_\phi(\mathbf{x}^l))$$

Ensemble the Scorers from Beams

- How to combine the scores of programs from K beam searches?
- The score of program **a** under context **x** given base models $\Phi = \{\phi^k\}_{k=1}^K$

$$v_{\omega,\Phi}(\mathbf{a};\mathbf{x}) = \sum_k [v'_\omega(\mathbf{a};\mathbf{x}, s_\phi^k(\mathbf{x})) - \bar{v}_\omega(\mathbf{x})].$$

where $\bar{v}_\omega(\mathbf{x})$ is the average score for programs in beam $s_\phi^k(\mathbf{x})$

$$\bar{v}_\omega(\mathbf{x}) = \frac{1}{|s_\phi^k(\mathbf{x})|} \sum_{\mathbf{a} \in s_\phi^k(\mathbf{x})} v_\omega(\mathbf{a};\mathbf{x}, s_\phi^k(\mathbf{x}))$$

and v'_ω backs-off v_ω to $\bar{v}_\omega(\mathbf{x})$ whenever **a** is not in beam

$$v'_\omega(\mathbf{a};\mathbf{x}, s_\phi^k(\mathbf{x})) = \begin{cases} v_\omega(\mathbf{a};\mathbf{x}, s_\phi^k(\mathbf{x})), & \text{if } \mathbf{a} \in s_\phi^k(\mathbf{x}) \\ \bar{v}_\omega(\mathbf{x}), & \text{else} \end{cases}$$

[Biloki+ 2019]

Results

- The impact of NPP, stacked learning (LOO) and ensemble

Setting	Model	Dev (std)	$\Delta^\dagger$	Test (std)	$\Delta^\dagger$
Mean of MAPOs trained on a single train/dev split	MAPO	41.9(0.3)	-	43.1(0.5)	-
	MAPO + NPP	42.4(0.7)	0.8	43.7(0.6)	0.5
Mean of MAPOs trained on LOO splits	MAPO	41.7(1.1)	-	42.8(0.5)	-
	MAPO + NPP*	43.0(0.2)$^+$	1.3	43.9(0.2)	1.1
Ensemble of 5 MAPOs trained on LOO splits	MAPO	-	-	45.5	-
	MAPO + NPP*	-	-	46.6	1.1
Ensemble of 10 MAPOs trained on LOO splits	MAPO	-	-	46.3(−)	-
	MAPO + NPP*	-	-	47.2(−)	0.9

Table 4: Main results. †Improvements compared to MAPO. *Stacked learning with Leave-One-Out (LOO) data splits. $^+$NPP uses 67%-33% train-dev splits from the stacked learning data.

Plan

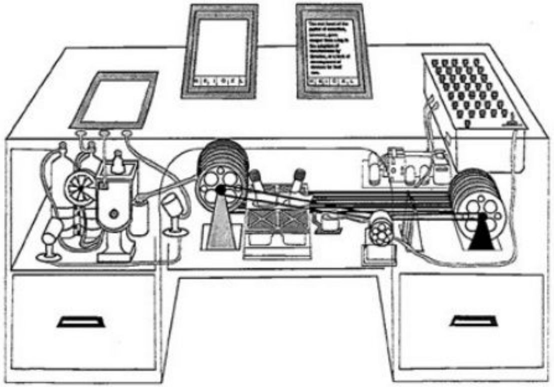
- ***Query understanding***
  - Weak supervision semantic parsing tasks
  - Neural Symbolic Machines
    - Symbolic representations for efficient inference
  - Memory Augmented Policy Optimization
    - Unbiased low-variance gradient estimation with experience replays
    - RL vs MML vs ML
  - Reranking for RL trained decoders
    - Sequence scorer, Stacked Learning, Scorer Ensembles
- ***Document understanding***
  - Generalizable, yet accountable & scalable
  - The return of inverted lists
  - Experiment with n-gram machines

[Vannevar Bush 1945]

A shared external memory

- The progress of civilizations depends on their shared memories

"AS WE MAY THINK"
(1945)

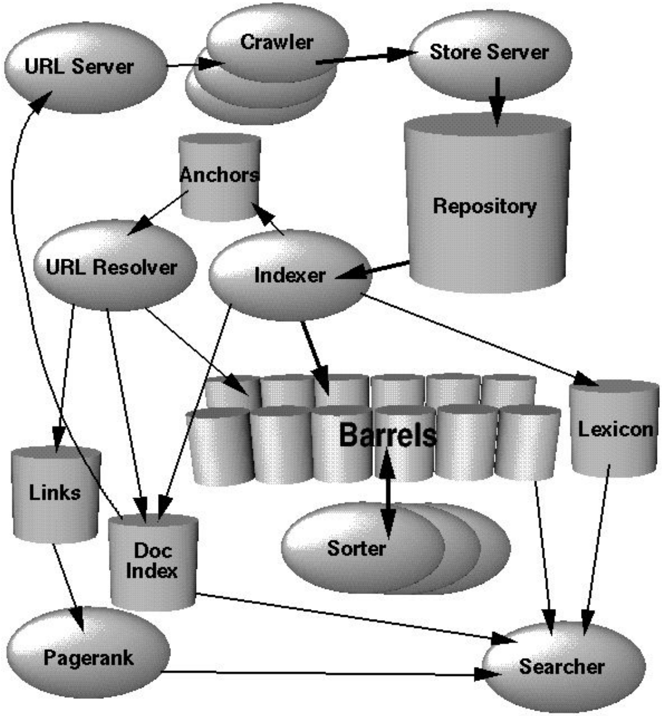


"Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, memex will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted with exceeding speed and flexibility. It is an enlarged intimate supplement to his memory.

[Brin & Page 1998]

Modern search engines

- Scalable
  - Can respond to user's requests within a fraction of a second
- Accountable
  - Reference, authority, engagement
- but less generalizable
  - work very well for popular queries through associations with engagements
  - distillation helps to the extent of available unlabeled queries
  - less generalizable to novel situations



Generalizable, yet accountable & scalable

- Generalizable
    - personal/contextual
    - compositional
    - cross lingual
    - multi-modality

- Accountable & Controllable
    - understand why a system produces certain result and promote trust
    - control model behavior through the update of its data

- Scalable
    - respond to user's requests within a fraction of a second

Transfer learning is promising for generalization

- ImageNet has thousands of labeled examples per class, while

- Children often only need one training example to learn a new concept

[Hernandez+ 2021]

Transfer learning favors big models

- the effective data transferred grows with the **model size** N and **fine-tuning data size** D_F

$$D_T = \text{effective data transferred} = k(D_F)^{\alpha}(N)^{\beta}$$

- and, in low data region

$$D_E = D_F + D_T \approx D_T$$

	Transfer Coefficients		
Transfer from	k	α	β
Text $\Longrightarrow$ Python	1.9e4	0.18	0.38
50% Text and 50% non-python code $\Longrightarrow$ Python	2.1e5	0.096	0.38

Lack of reasoning, accountability and efficiency

Q & A	Fact / Fiction
Q: what is barack obama's birthday? **A**: Barack Obama's birthday is August 4, 1961.	Fact
Q: what is barack obama's wife's birthday? **A**: Michelle Obama's birthday is January 17, 1964.	Fact
Q: what is barack obama's father's birthday? **A**: Barack Obama's father's birthday is August 4, 1961.	Fiction
Q: who are barack obama's daughters **A**: Sasha and Malia	Fact
Q: what are barack obama's daughters birthday? **A**: barack obama's daughters birthday is on august 4	Fiction
Q: what are Sasha and Malia obama's birthdays **A**: Sasha and Malia Obama's birthdays are on the same day, July 4, 1981.	Fiction

- Tested birthday fact questions on GPT-3 (2020, davinci)
- Given its source an answer could have been verified by the users or a cross encoder (2020 EfficientQA Competition)

Barack Hussein Obama Sr.,
Born: June 18th, 1934
https://en.wikipedia.org/wiki/Barack_Obama_Sr.

Sasha Obama, Born: June 10, 2001
Malia Ann Obama, Born: July 4, 1998
https://en.wikipedia.org/wiki/Family_of_Barack_Obama

[Garcia+ 1966]
[Wickman 2012]
[Bartol+ 2015]

Mammalian memory

- Very rapid adaptation (in just one or few trials) is necessary for survival
  - E.g., associating smell of food to sickness
- However, good representation is learnt gradually
  - e.g., learning during sleeps to prevent interference with established associations

Complementary Learning Theory

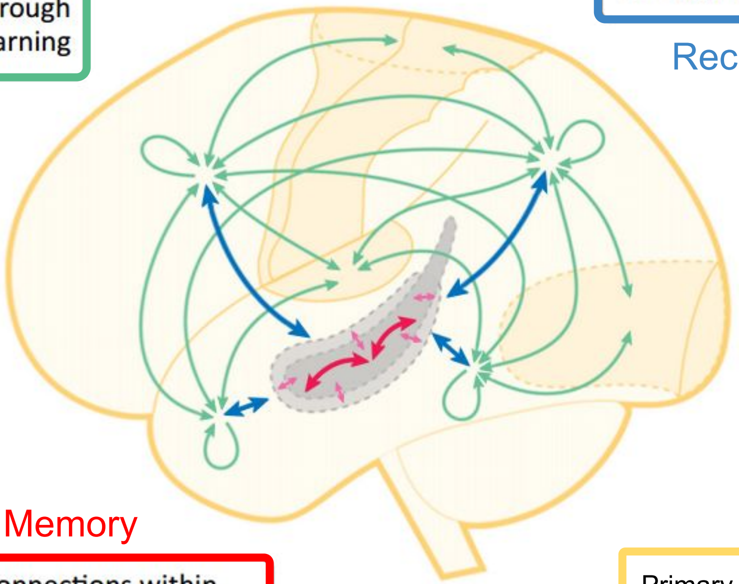
[McClelland+ 1995]
[Kumaran+ 2016]

Connections within and among neocortical areas (green) support gradual acquisition of structured knowledge through interleaved learning

Encode

Bidirectional connections (blue) link neocortical representations to the hippocampus/MTL for storage, retrieval, and replay

Record & Replay



Episodic Memory

Rapid learning in connections within hippocampus (red) supports initial learning of arbitrary new information

Observations

Primary sensory and motor cortices

[Khandelwal+ 2020]
[Guu+ 2020]
[Lewis+ 2020]

Explicit memory and control with retrieval-based models

- Generalizable
  - achieve generalizability with a small model + a big memory
- Accountable & Controllable
  - understand from which piece of memory certain result is deduced
  - the memory can be updated independent of the model
- Scalable
  - achieve low latency with an index structure to the memory

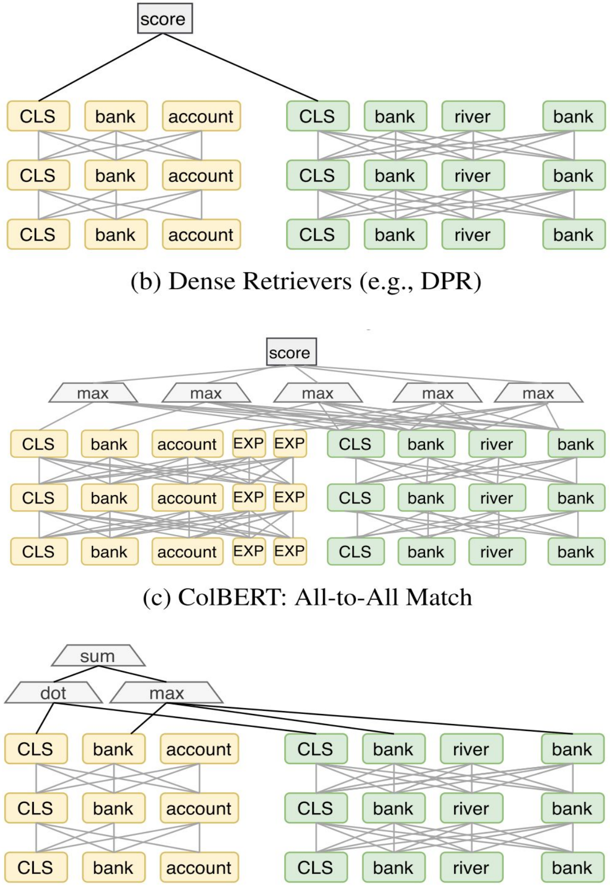
[Gao+ 2021]
[Lee+ 2021]

The return of inverted lists

- Classical IR systems
  - rely on exact lexical matches, which can carry out search efficiently with inverted list index
  - fall short of matching related terms (**vocabulary mismatch**) or modeling context of the terms (**semantic mismatch**).

- Dense Retrievers
  - lack of lexical matches
  - **huge indices** (100x) and **large latencies** (10x) especially for multi-vector representations like ColBERT, DensePhrase

- The two can be combined to get the best of both worlds



(b) Dense Retrievers (e.g., DPR)

(c) ColBERT: All-to-All Match

(d) COIL: Contextualized Exact Match

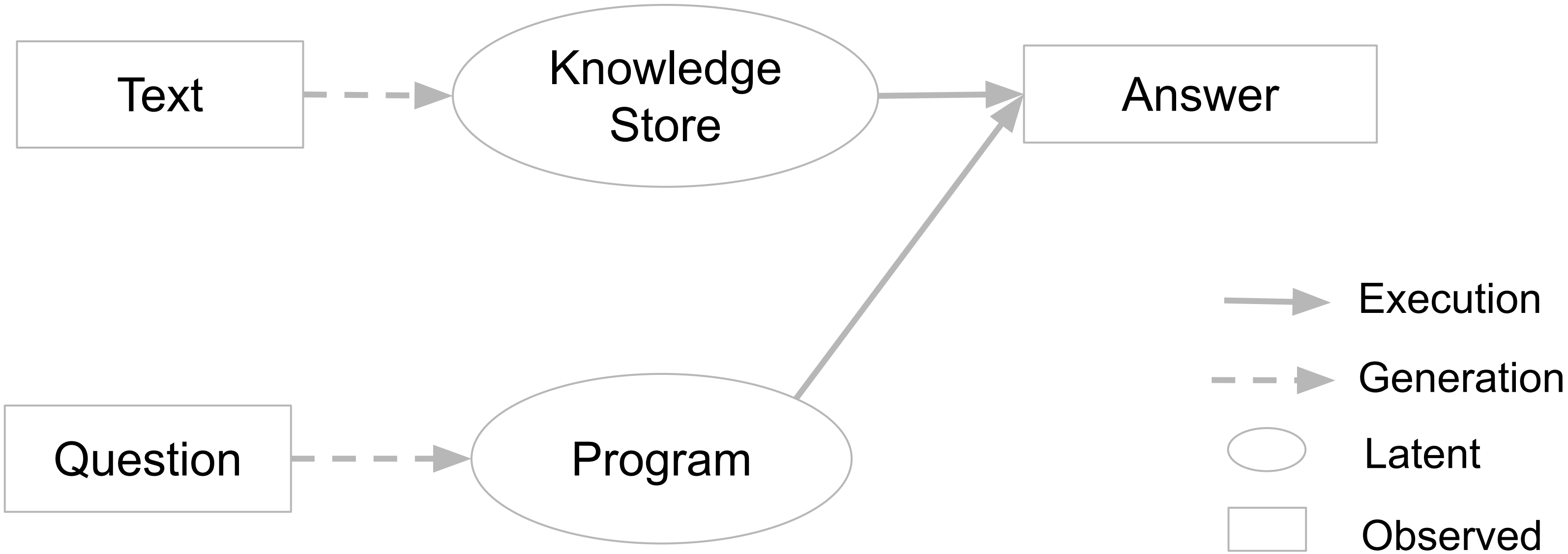
[Khandelwal+ 2020]
[Guu+ 2020]
[Lewis+ 2020]

Related work

- kNN-LM (Khandelwal+ 2020)
  - 2.9 perplexity improvement simply by linearly interpolating an LM's token prediction with the next token counts of k-nearest neighbors in the decoder state on training data
  - too expensive to perform retrieval during training
- REALM (Guu+ 2020), RAG (Lewis+ 2020)
  - augment sequence models with a latent knowledge (document) retriever for better interpretability and controllability
  - e2e training by backpropagating through a **dense doc retrieval** step
- This presentation
  - can we use a **single explicit knowledge representation**, i.e., short sequences of words, and achieve both **efficiency and semantic matching**?

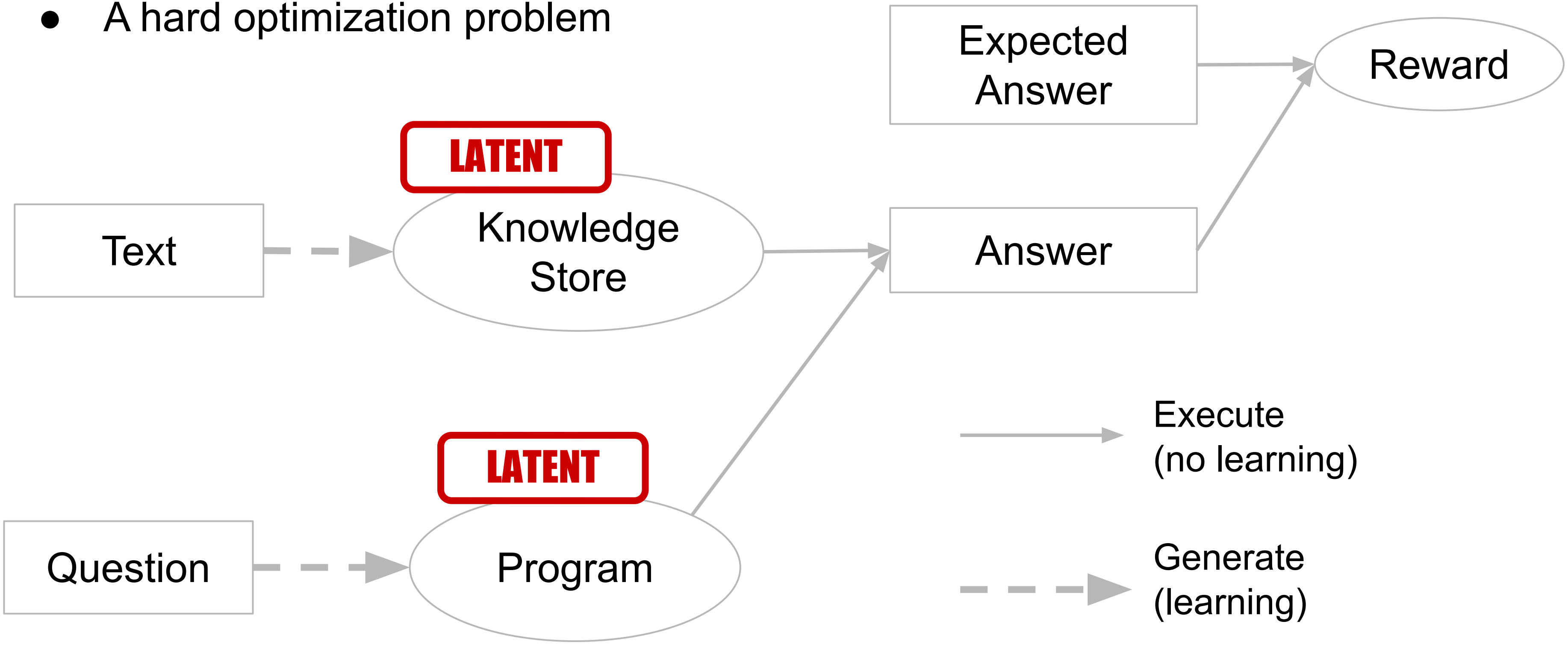
Question answering as a simple test bed

- A good representation should also support reasoning & scalability



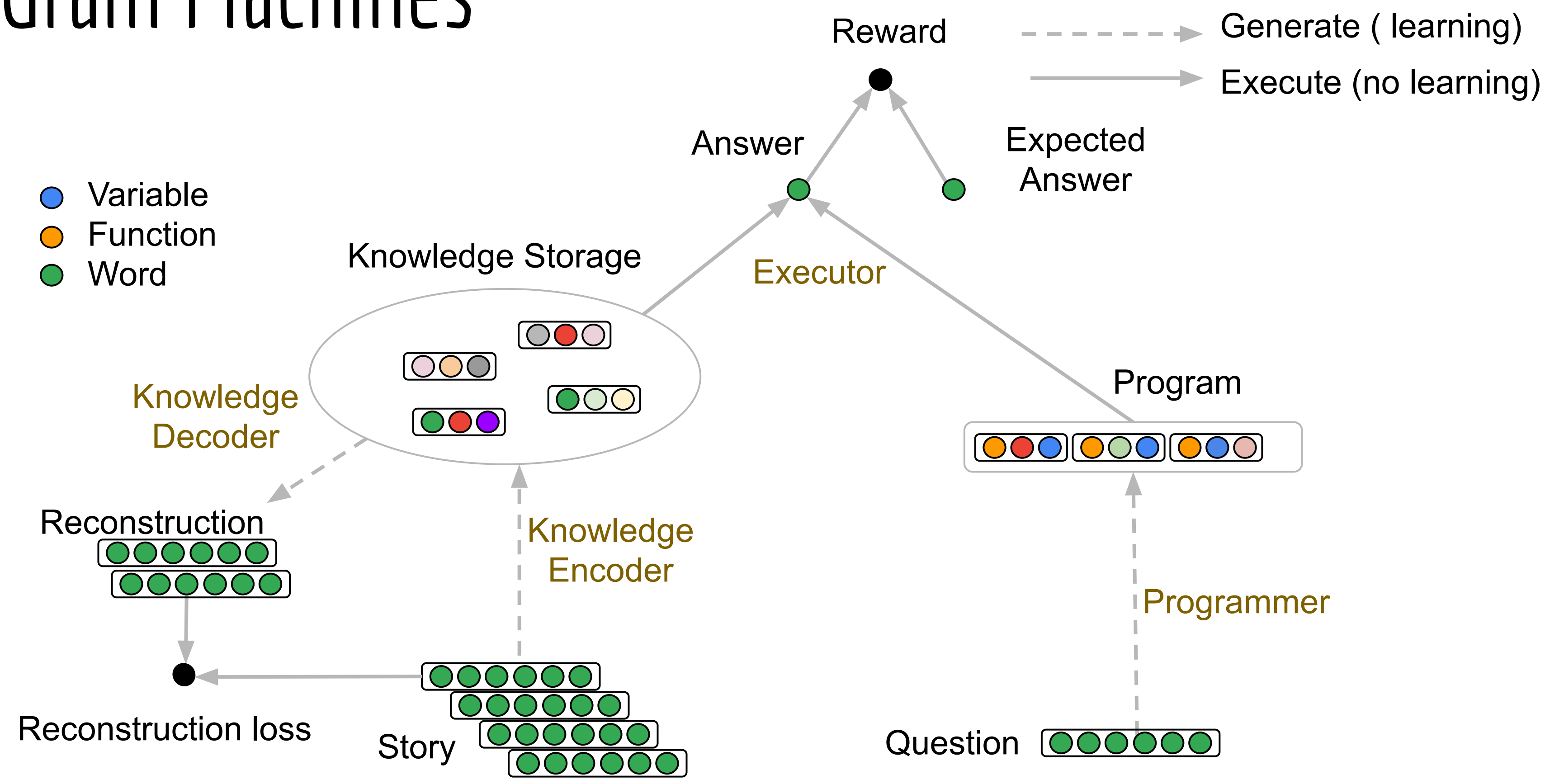
End-to-End Question Answering

- A hard optimization problem



N-Gram Machines

[Yang+ 2018]



Preliminary

[Yang+ 2018]

Seq2Seq components

- A **knowledge encoder** defines a distribution over knowledge tuples given sentences, and the distribution over knowledge store

$$P(\Gamma_i|s_i, s_{i-1}; \theta_{\text{enc}})$$

$$P(\mathbf{\Gamma}|\mathbf{s}; \theta_{\text{enc}}) = \Pi_{\Gamma_i \in \mathbf{\Gamma}} P(\Gamma_i|s_i, s_{i-1}; \theta_{\text{enc}})$$

- A **knowledge decoder** defines a distribution over sentences given tuples

$$P(s_i|\Gamma_i, s_{i-1}; \theta_{\text{dec}})$$

- A **programmer** defines a distribution over programs given a question

$$P(C|q, \mathbf{\Gamma}; \theta_{\text{prog}})$$

Preliminary

[Yang+ 2018]

Inference & Training

- Given an example (s, q, a)
  - maximize the **expected reward (QA) + sentence reconstruction (AE)**
    (VAE or contrastive loss might produce better latent space distributions)

$$O^{QA}(\theta_{\text{enc}}, \theta_{\text{prog}}) = \sum_{\mathbf{\Gamma}} \sum_{C} P(\mathbf{\Gamma}|\mathbf{s}; \theta_{\text{enc}}) P(C|q, \mathbf{\Gamma}; \theta_{\text{prog}}) R(\mathbf{\Gamma}, C, a),$$

$$O^{\text{AE}}(\theta_{\text{enc}}, \theta_{\text{dec}}) = \mathbb{E}_{p(z|x;\theta_{\text{enc}})}[\log p(x|z; \theta_{\text{dec}})] + \sum_{z \in \mathbf{Z}^N(x)} \log p(x|z; \theta_{\text{dec}}),$$

$Z^N(x)$: all tuples of length N which only consist of words from x

- Gradient estimation
  - **beam search** for its low variances

- Coordinate ascent
  - updates three components in alternation with **REINFORCE**

[Weston+ 2015]

Facebook bAbI Tasks

- Simulated question answering tasks to test the ability to "**understand**"
- We introduce a special version ("life-long bAbI"), which has stories of up to 10 million sentences

Sam walks into the kitchen. Sam picks up an apple. Sam walks into the bedroom. Sam drops the apple. Q: Where is the apple? A. Bedroom	Brian is a lion. Julius is a lion. Julius is white. Bernhard is green. Q: What color is Brian? A. White	Mary journeyed to the den. Mary went back to the kitchen. John journeyed to the bedroom. Mary discarded the milk. Q: Where was the milk before the den? A. Hallway

[Yang+ 2018]

Example Knowledge Store & Program

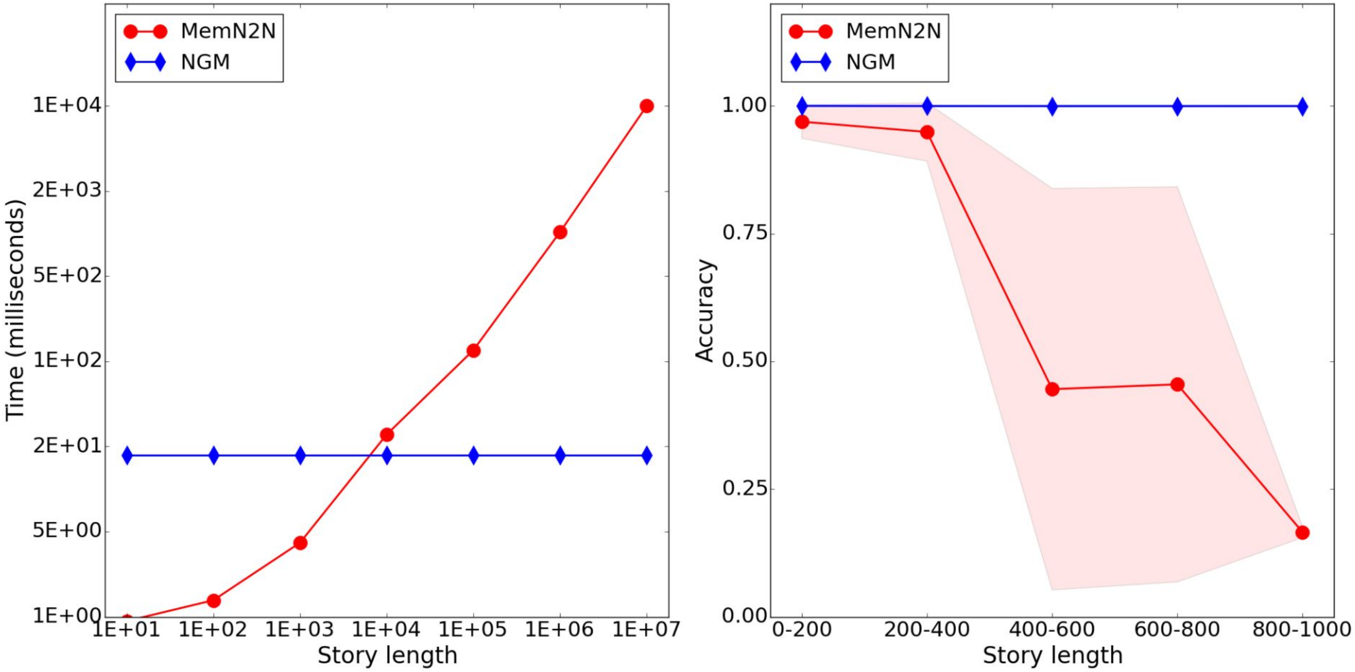
Table 6: Task 2 Two Supporting Facts

Story	Knowledge Storage
Sandra journeyed to the hallway. John journeyed to the bathroom. Sandra grabbed the football. Daniel travelled to the bedroom. John got the milk. John dropped the milk.	Sandra journeyed hallway John journeyed bathroom Sandra got football Daniel journeyed bedroom John got milk John got milk
Question	**Program**
Where is the milk?	ArgmaxFR milk got Argmax V1 journeyed

[Yang+ 2018]

Scalability

- A lot more scalable than commonly used deep model



Thanks!

Ranking models on MS MARCO (Tesla V100)

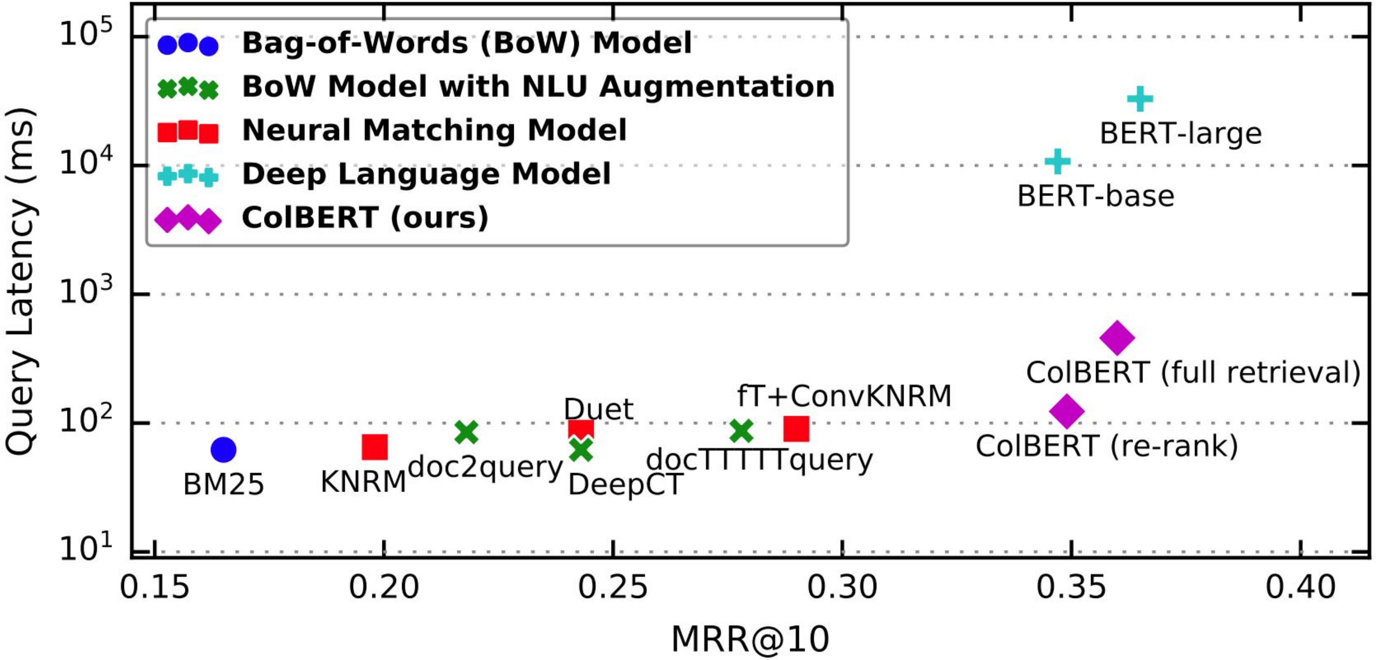
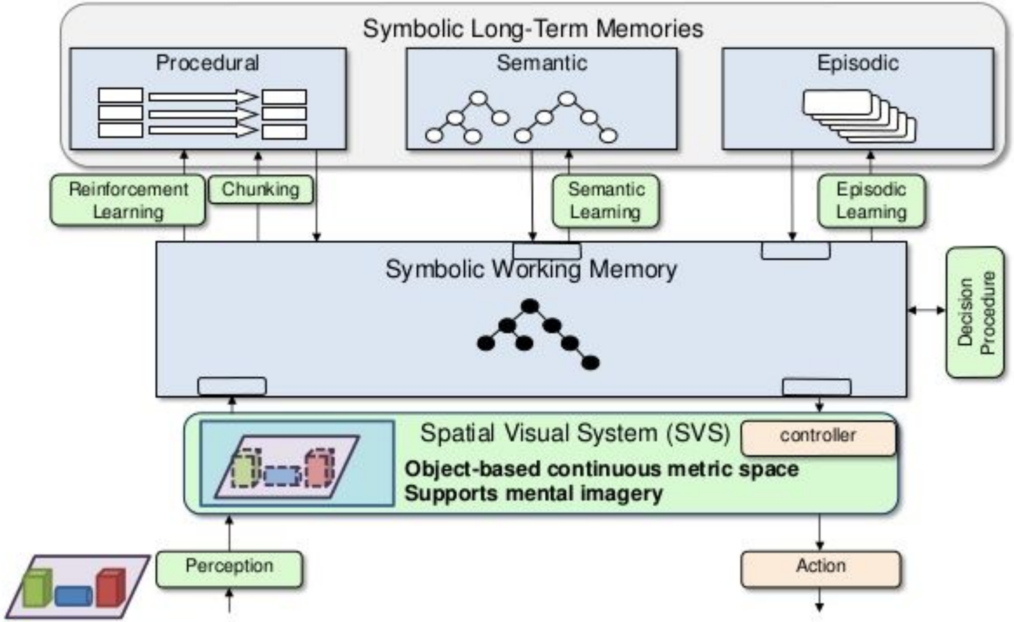


Table 3: Performance and latency of COIL systems with different representation dimensions. Results not applicable are denoted '–' and no available 'n.a.'. Here n_c denotes COIL CLS dimension and n_t token vector dimension. *: ColBERT use approximate search and quantization. We exclude I/O time from measurements.

Model		Dev Retrieval		DL2019 Retrieval		Latency/ms	
		MRR@10	Recall@1K	NDCG@10	MRR	CPU	GPU
BM25		0.184	0.853	0.506	0.825	36	n.a.
Dense		0.304	0.932	0.635	0.898	293	32
ColBERT		0.360	0.968	n.a.	n.a.	458*	–
COIL							
n_c	n_t						
768	32	0.355	0.963	0.704	0.924	380	41
128	32	0.350	0.953	0.692	0.956	125	23
128	8	0.347	0.956	0.694	0.977	113	21
0	32	0.341	0.949	0.660	0.915	67	18
0	8	0.336	0.940	0.678	0.953	55	16

Typical Cognitive Architectures

- The design of mammalian brains is inspiring to NLP systems
  - Since they are solving similar problems
- The design has not changed much since 30 years ago
  - “We’ve totally solved it already”
    -- Nate Derbinsky, Northeastern U.
- Lack of real applications, but now
  - internet economy and data
  - computation and machine learning



https://soar.eecs.umich.edu

Optimization

- For training stability and tweaking, we augment the training objective with **experience replays**

$$\nabla_{\theta_{\text{dec}}} O'(\theta) = \sum_{s_i \in \mathbf{s}} \sum_{\Gamma_i} [\beta(\Gamma_i) + P(\Gamma_i | s_i, s_{i-1}; \theta_{\text{enc}})] \nabla_{\theta_{\text{dec}}} \log P(s_i | \Gamma, s_{i-1}; \theta_{\text{dec}}),$$

$\beta(\Gamma_i)$ is 1 if Γ_i only contains tokens from s_i and 0 otherwise

$$\nabla_{\theta_{\text{enc}}} O'(\theta) = \sum_{s_i \in \mathbf{s}} \sum_{\Gamma_i} [P(\Gamma_i | s_i, s_{i-1}; \theta_{\text{enc}}) \log P(s_i | \Gamma_i, s_{i-1}; \theta_{\text{dec}}) + \mathcal{R}(\mathcal{G}'(\Gamma_i)) + \mathcal{R}(\mathcal{G}(\Gamma_i))] \nabla_{\theta_{\text{enc}}} \log P(\Gamma_i | s_i, s_{i-1}; \theta_{\text{enc}}),$$

where $\mathcal{R}(\mathcal{G}) = \sum_{\mathbf{\Gamma} \in \mathcal{G}} \sum_C P(\mathbf{\Gamma} | \mathbf{s}; \theta_{\text{enc}}) P(C | q, \mathbf{\Gamma}; \theta_{\text{prog}}) R(\mathbf{\Gamma}, C, a)$ is the total expected reward for a set of valid knowledge stores $\mathcal{G}$, $\mathcal{G}(\Gamma_i)$ is the set of knowledge stores which contains the tuple Γ_i, and $\mathcal{G}'(\Gamma_i)$ is the set of knowledge stores which contains the tuple Γ_i through tweaking.

$$\nabla_{\theta_{\text{prog}}} O'(\theta) = \sum_{\mathbf{\Gamma}} \sum_{C} [\alpha I [C \in \mathcal{C}^*(\mathbf{s}, q)] + P(C | q, \mathbf{\Gamma}; \theta_{\text{prog}})] \cdot P(\mathbf{\Gamma} | \mathbf{s}; \theta_{\text{enc}}) R(\mathbf{\Gamma}, C, a) \nabla_{\theta_{\text{prog}}} \log P(C | q, \mathbf{\Gamma}; \theta_{\text{prog}}),$$

where $\mathcal{C}^*(\mathbf{s}, q)$ is the experience replay buffer for $(\mathbf{s}, q)$. $\alpha = 0.1$ is a constant. During training, the program with the highest weighted reward (i.e. $P(\mathbf{\Gamma} | \mathbf{s}; \theta_{\text{enc}}) R(\mathbf{\Gamma}, C, a)$) is added to the replay buffer.

- optimize by **coordinate ascent** – updating three components in alternation with **REINFORCE**

Details

[Liang+, 2018]

WikiTableQuestions: example solutions

Superlative	
nt-13901: the most points were scored by which player?	
(argmax all_rows r.points-num)	Sort all rows by column 'points' and take the first row.
(hop v0 r.player-str)	Output the value of column 'player' for the rows in v0.
Difference	
nt-457: how many more passengers flew to los angeles than to saskatoon?	
(filter$_{in}$ all_rows ['saskatoon'] r.city-str)	Find the row with 'saskatoon' matched in column 'city'.
(filter$_{in}$ all_rows ['los angeles'] r.city-str)	Find the row with 'los angeles' matched in column 'city'.
(diff v1 v0 r.passengers-num)	Calculate the difference of the values in the column 'passenger' of v0 and v1.

Details

[Liang+, 2018]

More examples

Before / After	
nt-10832: which nation is before peru?	
(filter$_{in}$ all_rows [‘peru’] r.nation-str)	Find the row with ‘peru’ matched in ‘nation’ column.
(previous v0)	Find the row before v0.
(hop v1 r.nation-str)	Output the value of column ‘nation’ of v1.

Compare & Count	
nt-647: in how many games did sri lanka score at least 2 goals?	
(filter$_{\geq}$ all_rows [2] r.score-num)	Select the rows whose value in the ‘score’ column >= 2.
(count v0)	Count the number of rows in v0.

Exclusion	
nt-1133: other than william stuart price, which other businessman was born in tulsa?	
(filter$_{in}$ all_rows [‘tulsa’] r.hometown-str)	Find rows with ‘tulsa’ matched in column ‘hometown’.
(filter$_{!in}$ v0 [‘william stuart price’] r.name-str)	Drop rows with ‘william stuart price’ matched in the value of column ‘name’.
(hop v1 r.name-str)	Output the value of column ‘name’ of v1.

Details

[Liang+ 2017]

Results on WebQuestionsSP

- First end-to-end seq2seq to achieve SOTA on semantic parsing with weak supervision over large knowledge base
- The performance approached SOTA with full supervision

Model	Avg. Prec.@1	Avg. Rec.@1	Avg. F1@1
STAGG	67.3	73.1	66.8
NSM – our model	70.8	76.0	**69.0**
STAGG (full supervision)	70.9	80.3	71.7

Comparison of model update strategies

[Liang+ 2018]

Correctness of the behavior

Experiences & Reward

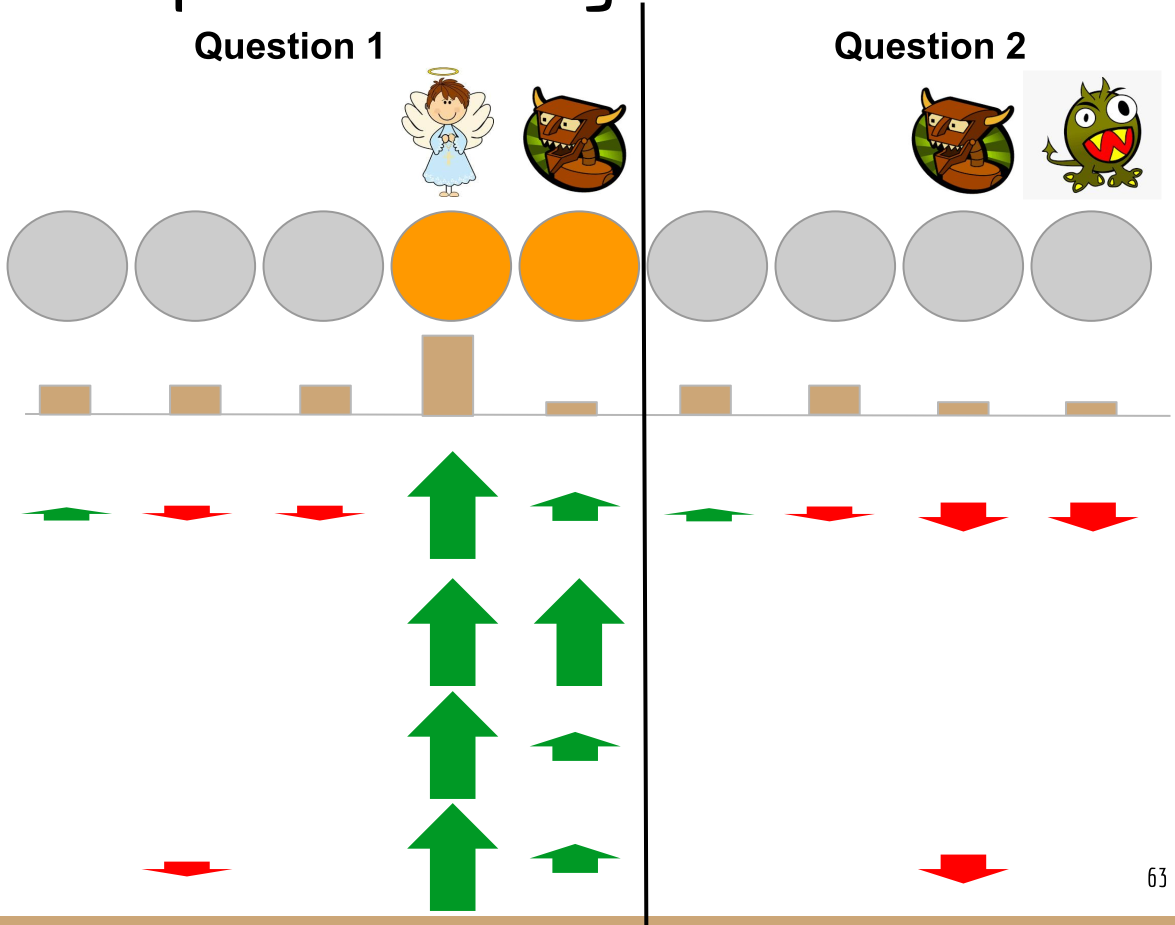
Model’s preference

On-policy optimization (REINFORCE)

Iterative Maximum Likelihood (IML)

Maximum Marginal Likelihood (MML)

MAPO

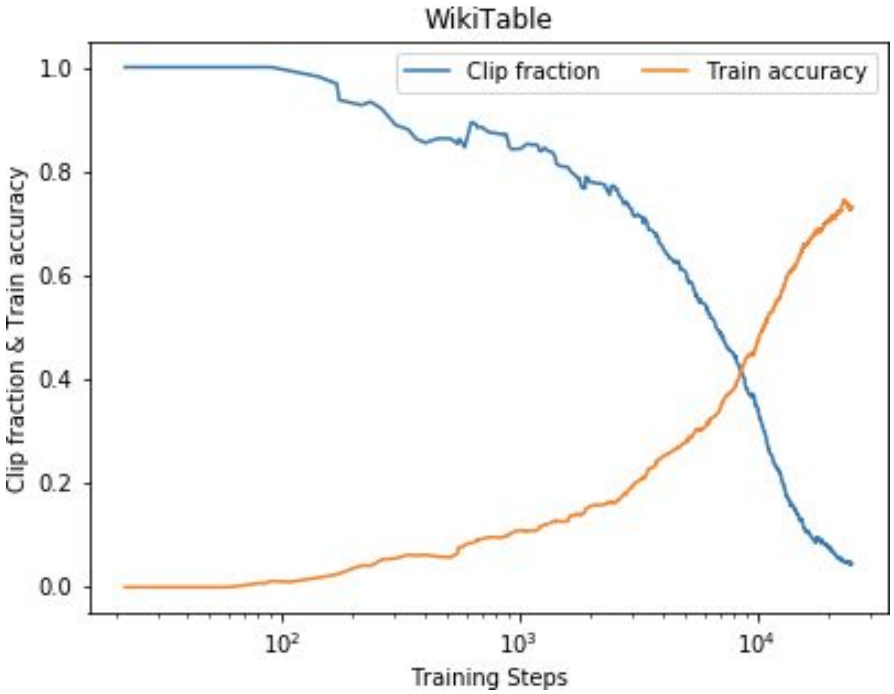


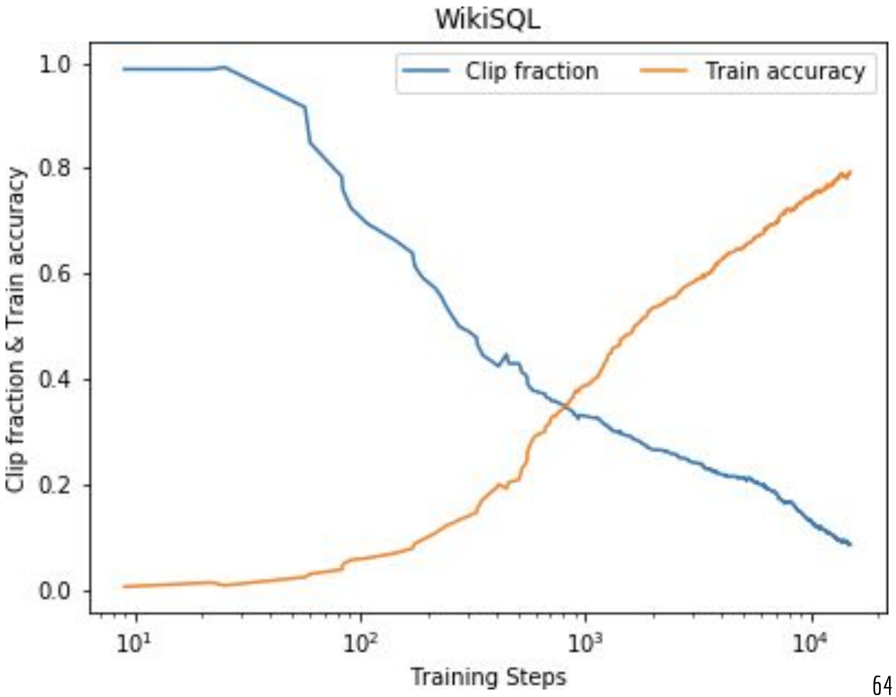
Details

[Liang+ 2018]

Clipping Mechanism

- Training becomes less biased over time





Details

[Liang+ 2018]

Results with weak supervision

Model	E.S.	Dev.	Test
Pasupat & Liang (2015)[28]	-	37.0	37.1
Neelakantan et al. (2017)[26]	1	34.1	34.2
Neelakantan et al. (2017)[26]	15	37.5	37.7
Haug et al. (2017)[15]	1	-	34.8
Haug et al. (2017)[15]	15	-	38.7
Zhang et al. (2017)[51]	-	40.4	43.7
MAPO	1	42.7	43.8
MAPO (ensembled)	5	-	46.2

Table 3: Results on WIKITABLEQUESTIONS. E.S. is the number of ensembles (if applicable).

Model	Dev.	Test
Zhong et al. (2017)[52]*	60.8	59.4
Wang et al. (2017)[40]*	67.1	66.8
Xu et al. (2017)[46]*	69.8	68.0
Huang et al. (2018)[18]*	68.3	68.0
Yu et al. (2018)[48]*	74.5	73.5
Sun et al. (2018)[38]*	75.1	74.6
Dong & Lapata (2018)[12]*	79.0	78.5
MAPO	72.4	72.6
MAPO (ensemble of 5)	-	74.9

Table 4: Results on WIKISQL. *All other methods use question-program pairs as strong supervision, while MAPO only uses question-answer pairs as weak supervision.