Weakly Supervised Natural Language Understanding

Ni Lao
mosaix.ai
1.3.2019

Plan

- ***Weak Supervision NLP***
  - NLP & software 2.0
  - Semantics as a foreign language
  - Semantic search
  - Semantic Parsing Tasks -- WebQuestionsSP, WikiTableQuestions
- ***Neural Symbolic Machines*** (ACL 2017)
  - Compositionality (short term memory)
  - Scalable KB inference (symbolism)
- ***Memory Augmented Policy Optimization*** (NIPS 2018)
  - Experience replay & optimal updating strategy
  - RL vs MML vs ML

Language understanding for AI and humanity

- Experts with different views of AI agree on the potential of NLP

If you got a billion dollars to spend on a huge research project that you get to lead, what would you like to do?
-- r/CyberByte, 2015

NLP is fascinating, allowing us to focus on **highly-structured inference** problems, on issues that go to the core of "**what is thought**" but remain eminently practical, and on a technology that surely would **make the world a better place**.
-- Michael I Jordan

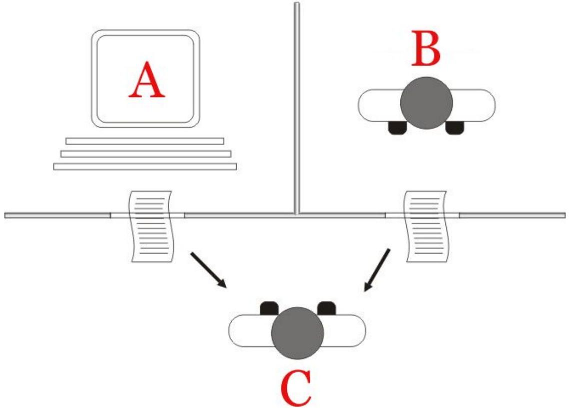
What kind of impact you hope deep learning has on our future?
-- Steve Paikin, 2016

I hope it allows Google to … search by the **content of the document** rather than by the words in the document … I hope it will make for **intelligent personal** systems, who can **answer questions** in a sensible way … It will make computers much easier to use. Because you'll be able to just **say to your computer** "print this damn thing"
-- Geoffrey Hinton

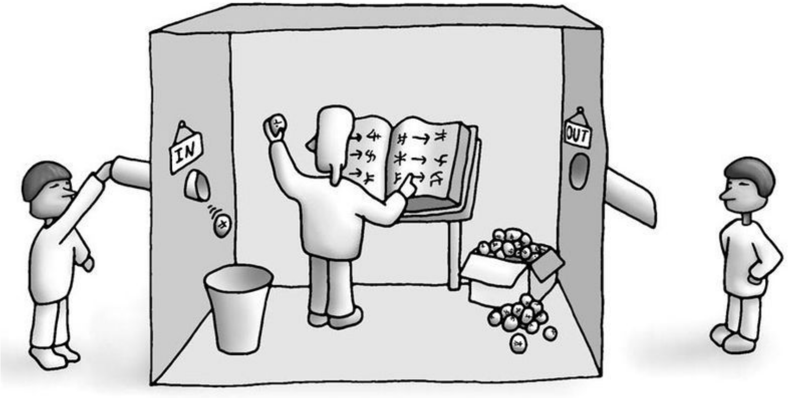
What is understanding?

"If they find a parrot who could **answer** to everything, I would claim it to be an **intelligent** being without hesitation.",
-- Alan Turing, 1950

Does the machine literally "**understand**" Chinese ? Or is it merely **simulating** the ability to understand Chinese?
-- John Searle, 1980



The Imitation Game



The Chinese Room Argument

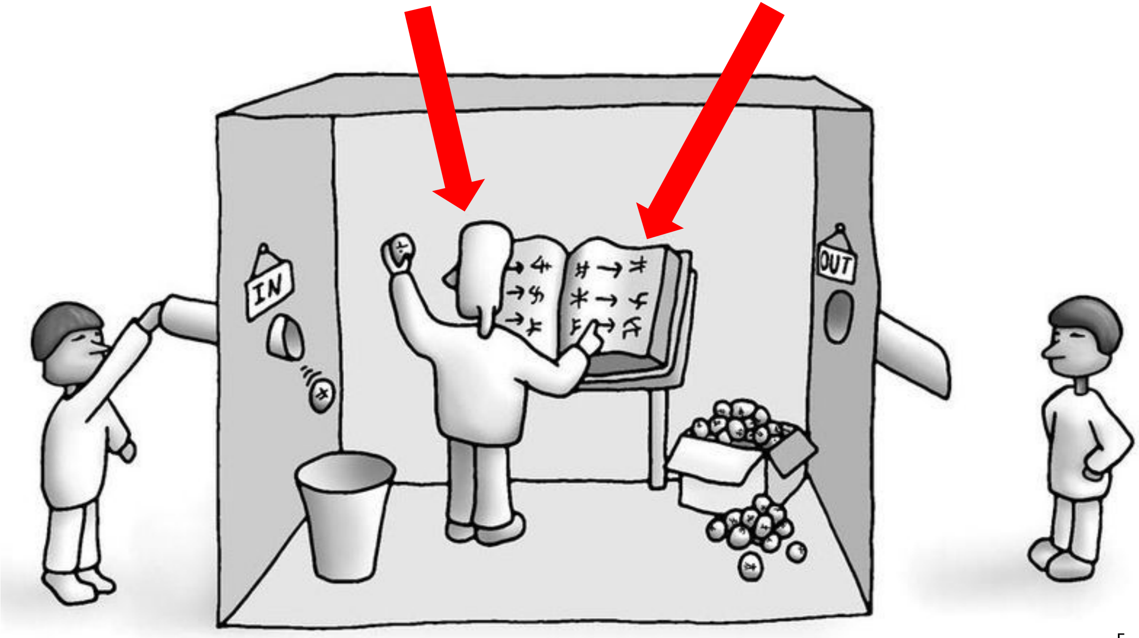
Full Supervision NLP

- Traditionally NLP is a labor-intensive business

Applications	
Discourse Processing	
Semantic Parsing	
Syntactic analysis	
Morphological analysis	
ASR	OCR
speech	text

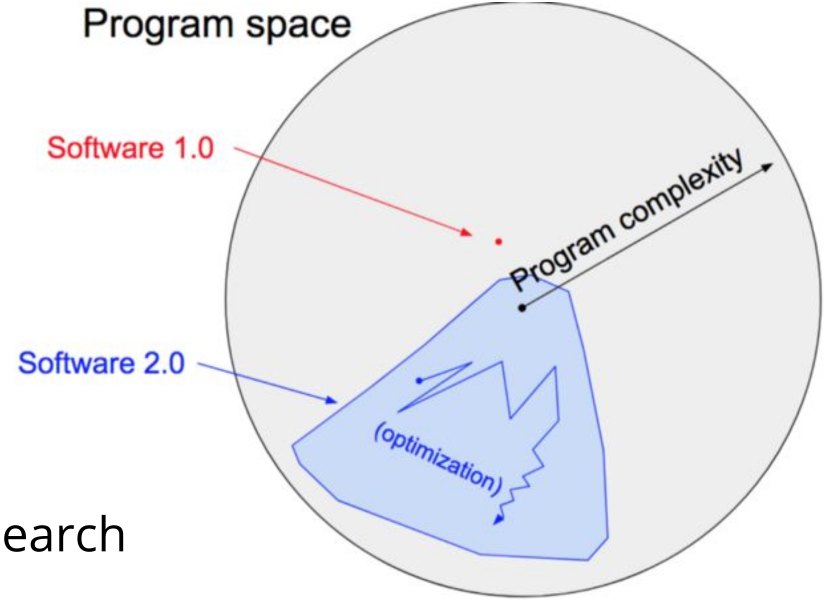
NLP algorithms simply lookup

Rules written by 1000 PhDs



Software 2.0

1. specify some goal on the behavior
   - e.g., “satisfy input output pairs of examples”,
   - e.g.,“win a game of Go”
2. write a rough skeleton of the code that identifies a subset of program space to search
   - e.g. a neural net architecture
3. use the computational resources at disposal to search this space for a program that works.



[Karpathy 2017; Watson 2017; Ratner+ 2018]

Death of feature engineering. (The) **users** of the software will (play) a direct role in building it. **Data labeling** is a central component to system design.

Where does knowledge come from?

- Can only cover the most popular semantics used by human

the world

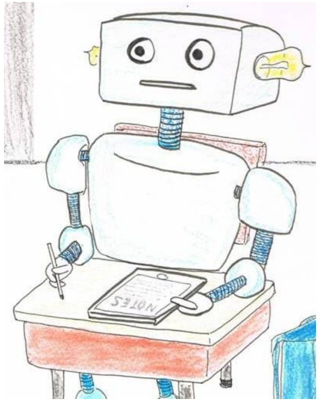
domain experts
(bottlenecks)

expert systems with
knowledge bases

Weak Supervision NLP

- Avoid the knowledge acquisition bottleneck with machine learning
- Then we can cover more semantics used by human

end to end examples
(e.g., QA pairs)

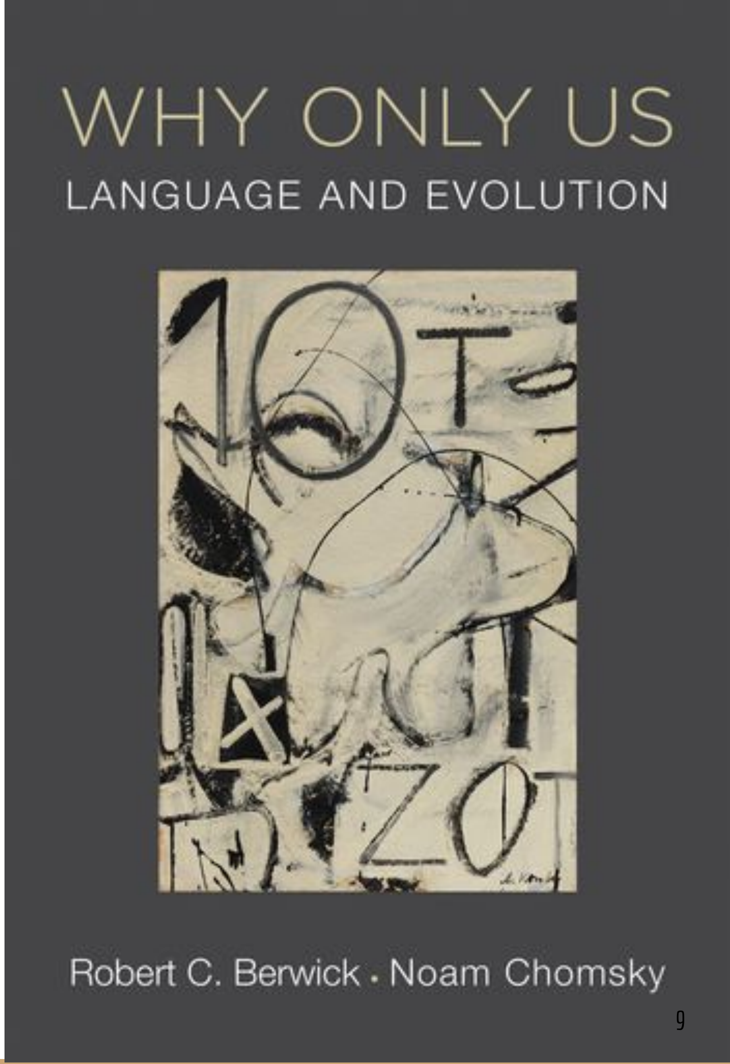


machine learning

intelligent systems
with knowledge

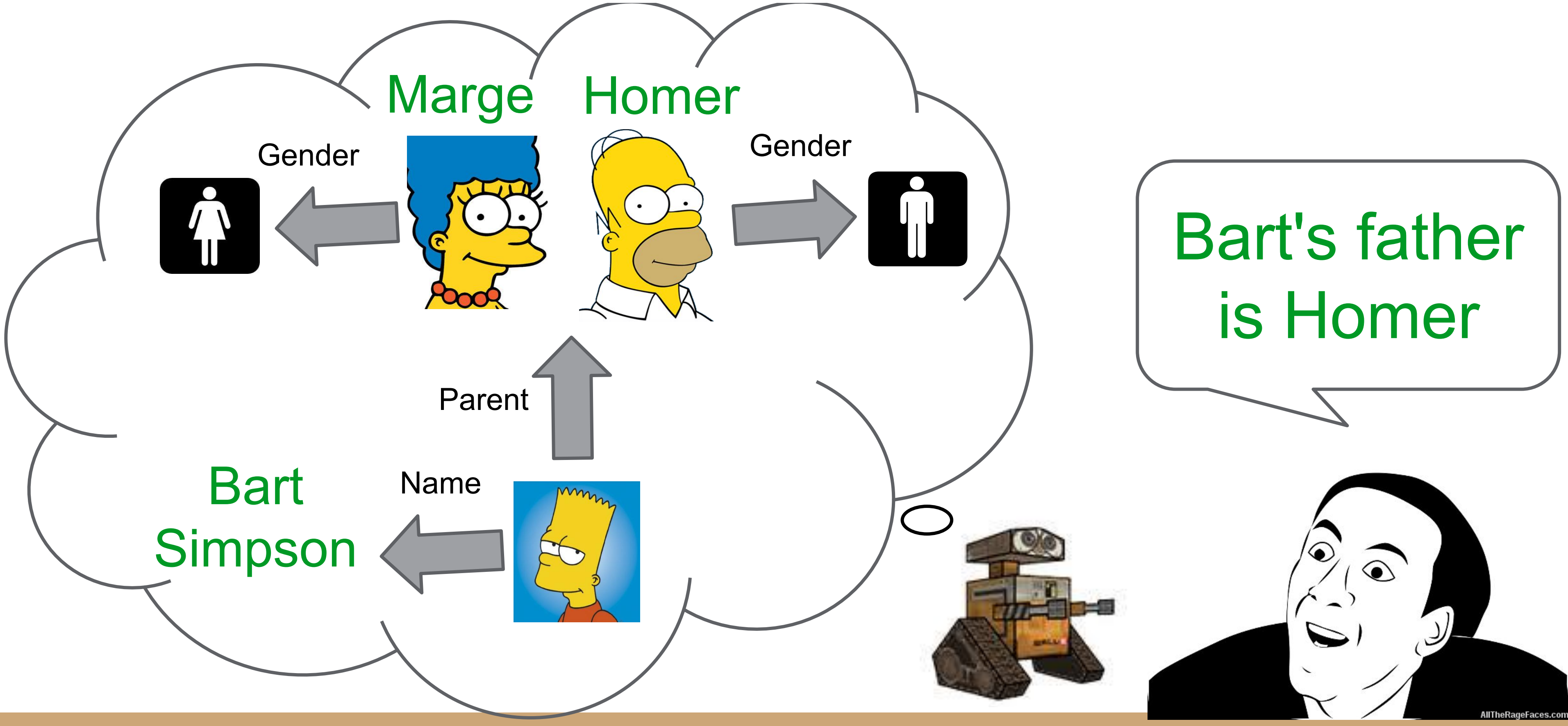
Language & Reasoning

- The formalist view
  - Language was primarily invented for reasoning [Everaert+ 2015]
- The functionalist view
  - Language is for communication [Kirby 2017]
- Cognitive coupling hypothesis
  - sequential processing is “necessary for behaviours such as primate tool use, navigation, foraging and social action.” [Kolodny & Edelman 2018]



[Liang+ 2017]

Reasoning is needed to understand text



Semantics is a language for 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]

Semantics is a language for computation

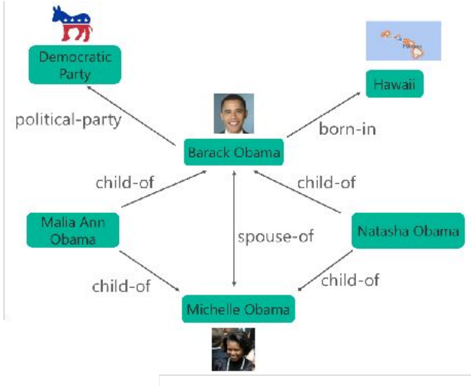
Semantic parsing

Largest city in US?

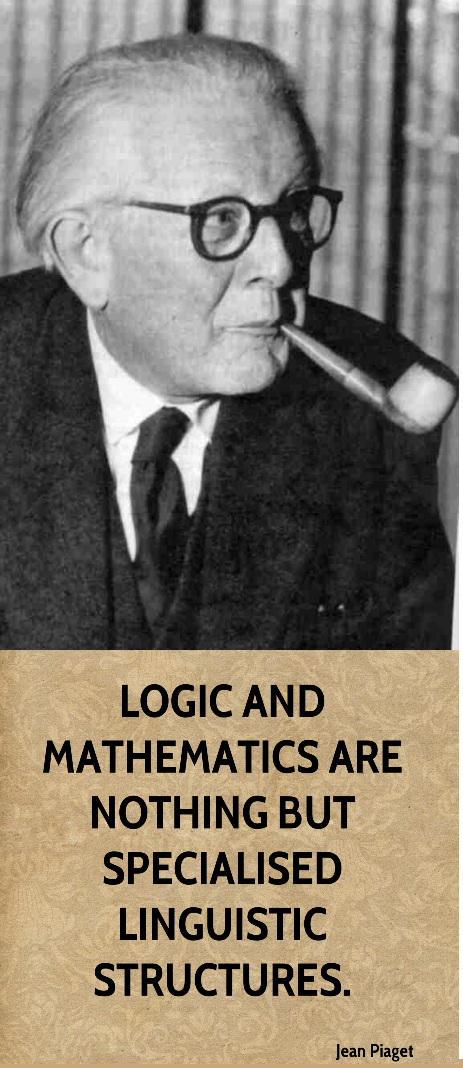
```
GO
(Hop V1 CityIn)
(Argmax V2 Population)
RETURN
```

Computation

NYC



Freebase, DBpedia, YAGO, NELL



Semantics as a foreign language

1) Natural languages are programming languages to control human behavior (either others or self)

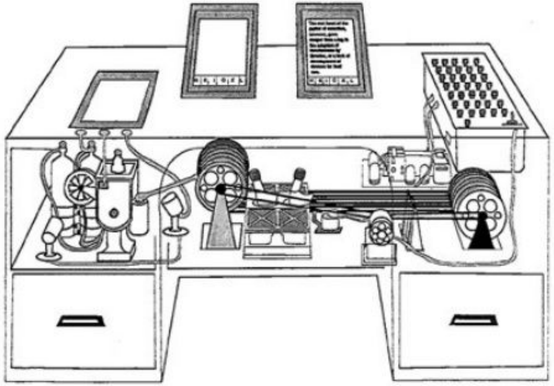
2) For machines and human to understand each other, they just need translation models trained with control theory

[Vannevar Bush 1945]

Internet as an external memory

- How information should be organized for scalability?

"AS WE MAY THINK"
(1945)

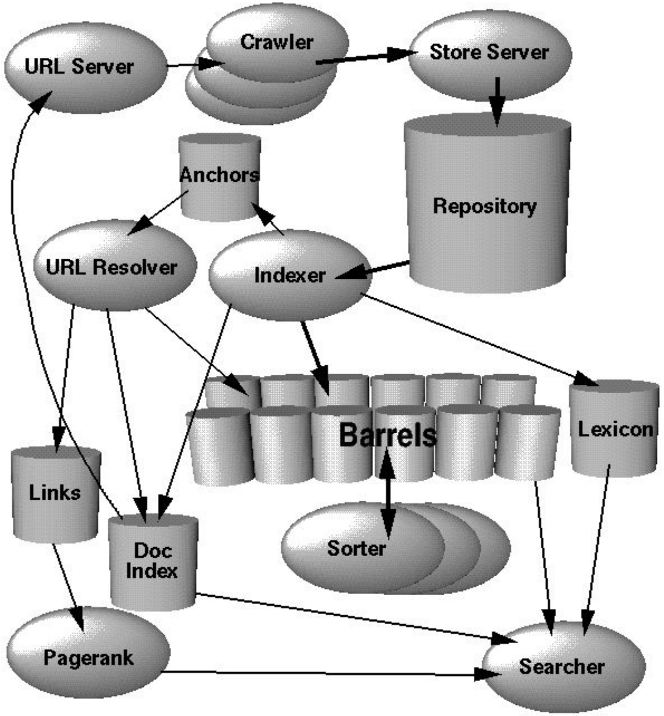


"Consider a future device for individual use, which is a sort of mechanized private file and library. It needs a name, and to coin one at random, memex will do. A memex is a device in which an individual stores all his books, records, and communications, and which is mechanized so that it may be consulted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

[Brin & Page 1998]

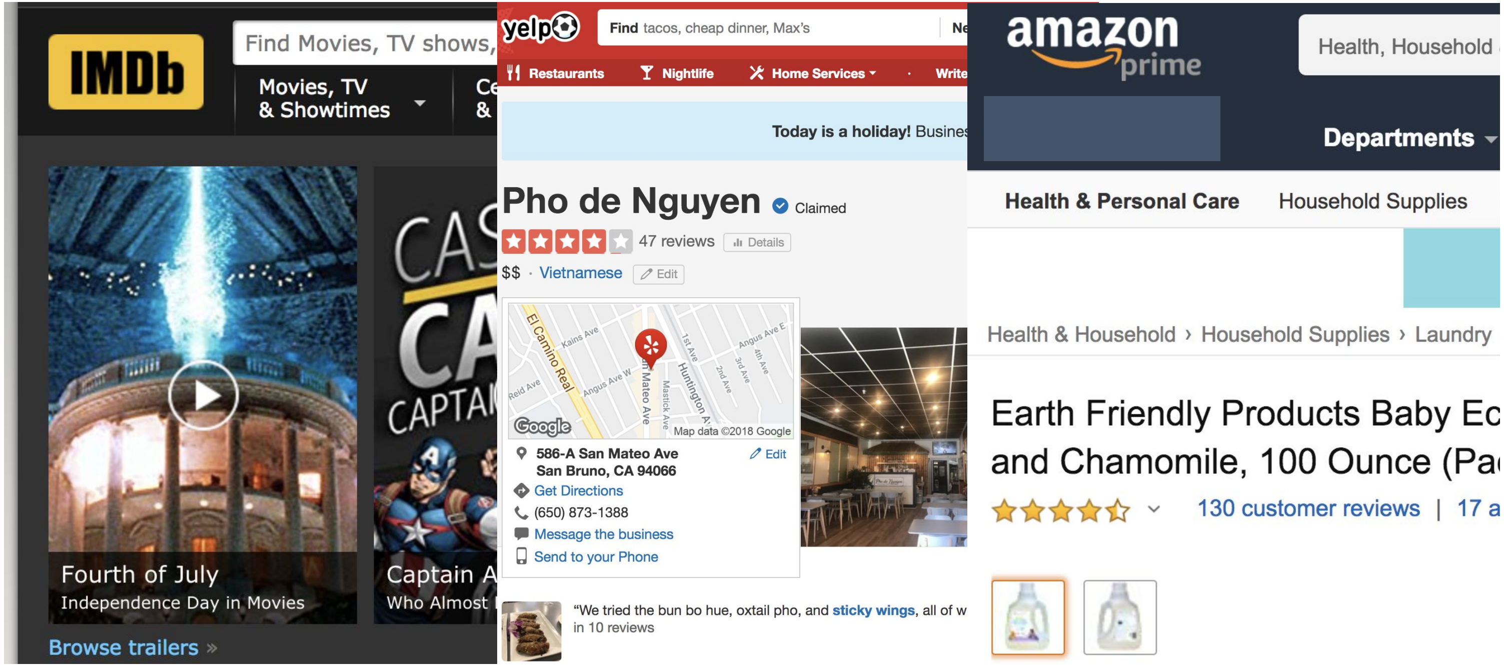
The scalability of modern search engines

- Can respond to user's requests within a fraction of a second
- But are weak at text understanding and complex reasoning



Can we search entities on the web?

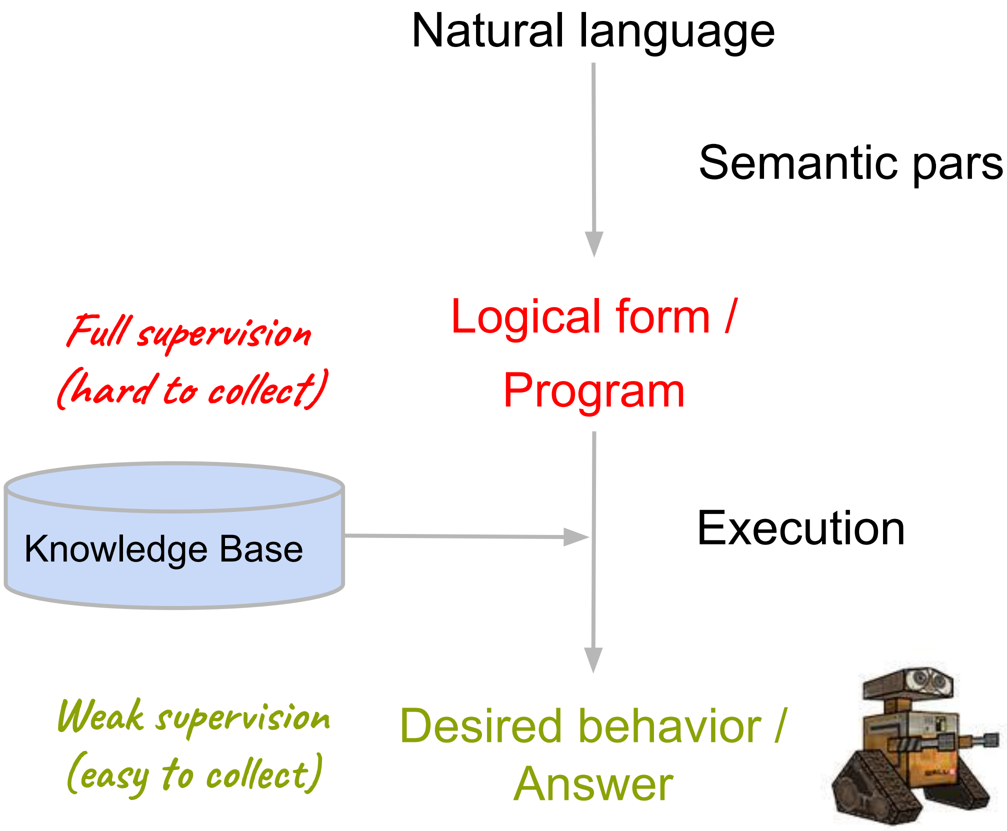
- Multimedia, business, products have a lot of reviews and descriptions



[Berant+ 2013]
[Liang 2013]

Semantic Parsing

- Natural language queries or commands are converted to computation steps on data and produce the expected answers or behavior



Details

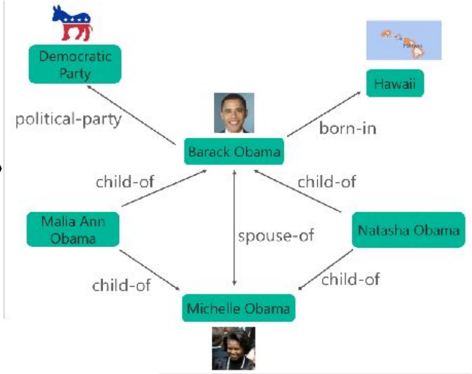
Related Works

- Training from full supervision is labor-intensive
  - DeepCoder [Balog, 2016]
  - NPI [Reed & Freitas, 2015]
  - Seq2Tree [Dong & Lapata, 2016]
  - …
- Traditional semantic parsing models require feature engineering
  - SEMPRE [Berant et al, 2013]
  - STAGG [Yih et al, 2015]
  - …
- End-to-end differentiable models cannot scale to large databases
  - Neural Programmer [Neekalantan et al, 2015]
  - Neural Turing Machines [Graves et al, 2014]
  - Neural GPU [Kaiser & Sutskever, 2015]
  - …
- Combine deep learning, symbolic reasoning and reinforcement learning
  - Neural Symbolic Machines (Liang, Berant, Le, Forbus, Lao, 2017)
  - Memory Augmented Policy Optimization (MAPO) (Liang, Norouzi, Berant, Le, Lao, 2018)

Question Answering with Knowledge Base

Largest city in US?

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



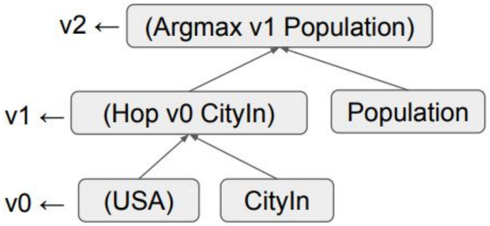
Freebase, DBpedia, YAGO, NELL

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 (Optimization)

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

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

WebQuestionsSP Dataset

- 5,810 questions from Google Suggest API & Amazon MTurk[1]
- Remove invalid QA pairs[2]
- 3,098 training examples, 1,639 testing examples remaining
- Open-domain and contains grammatical error
- Multiple entities as answer => macro-averaged F1

Grammatical error	**Multiple entities**
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

Details

Brief Summary & Preview

DL leads to the death of feature engineering (but not domain knowledge)
DL makes it easier to leverage pre-trained unsupervised models

	SEMPRE/STAGG	**This talk**
Paraphrase (semi-supervised)	Co-occurrences collected from 1B docs (ClueWeb) and Freebase	Embeddings trained from 840B text tokens (GloVe)
Compositionality of semantics	Relies on syntactic structure (text spans) and domain specific rules to constrain the generation of logic forms	A (LISP like) language specifies computations on KG
Search in the program space	Priority queue & heuristic rules	RL & a program interpreter with syntactical & semantic checks
Modeling	Expert designed features	Deep learning

[Pasupat & Liang, 2015]

WikiTableQuestions: Dataset

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: $\{10\}$

[Pasupat & Liang, 2015]
[Liang+ 2018]

WikiTableQuestions: semantics

Function	Arguments	Returns	Description
(**hop** v p)	**v**: a list of rows. **p**: a column.	a list of cells.	Select the given column of the given rows.
(**argmax** v p) (**argmin** v p)	**v**: a list of rows. **p**: a number or date column.	a list of rows.	From the given rows, select the ones with the largest / smallest value in the given column.
(**filter**$_{>}$ v q p) (**filter**$_{\geqslant}$ v q p) (**filter**$_{<}$ v q p) (**filter**$_{\leqslant}$ v q p) (**filter**$_{=}$ v q p) (**filter**$_{\neq}$ v q p)	**v**: a list of rows. **q**: a number or date. **p**: a number or date column.	a list of rows.	From the given rows, select the ones whose given column has certain order relation with the given value.
(**filter**$_{in}$ v q p) (**filter**$_{!in}$ v q p)	**v**: a list of rows. **q**: a string. **p**: a string column.	a list of rows.	From the given rows, select the ones whose given column contain / do not contain the given string.

[Pasupat & Liang, 2015]
[Liang+ 2018]

WikiTableQuestions: semantics

(**first** v) (**last** v)	**v**: a list of rows.	a row.	From the given rows, select the one with the smallest / largest index.
(**previous** v) (**next** v)	**v**: a row.	a row.	Select the row that is above / below the given row.
(**count** v)	**v**: a list of rows.	a number.	Count the number of given rows.
(**max** v p) (**min** v p) (**average** v p) (**sum** v p)	**v**: a list of rows. **p**: a number column.	a number.	Compute the maximum / minimum / average / sum of the given column in the given rows.
(**mode** v p)	**v**: a list of rows. **p**: a column.	a cell.	Get the most common value of the given column in the given rows.
* (**same_as** v p)	**v**: a row. **p**: a column.	a list of rows.	Get the rows whose given column is the same as the given row.
(**diff** v0 v1 p)	**v0**: a row. **v1**: a row. **p**: a number column.	a number.	Compute the difference in the given column of the given two rows.

[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.

[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.

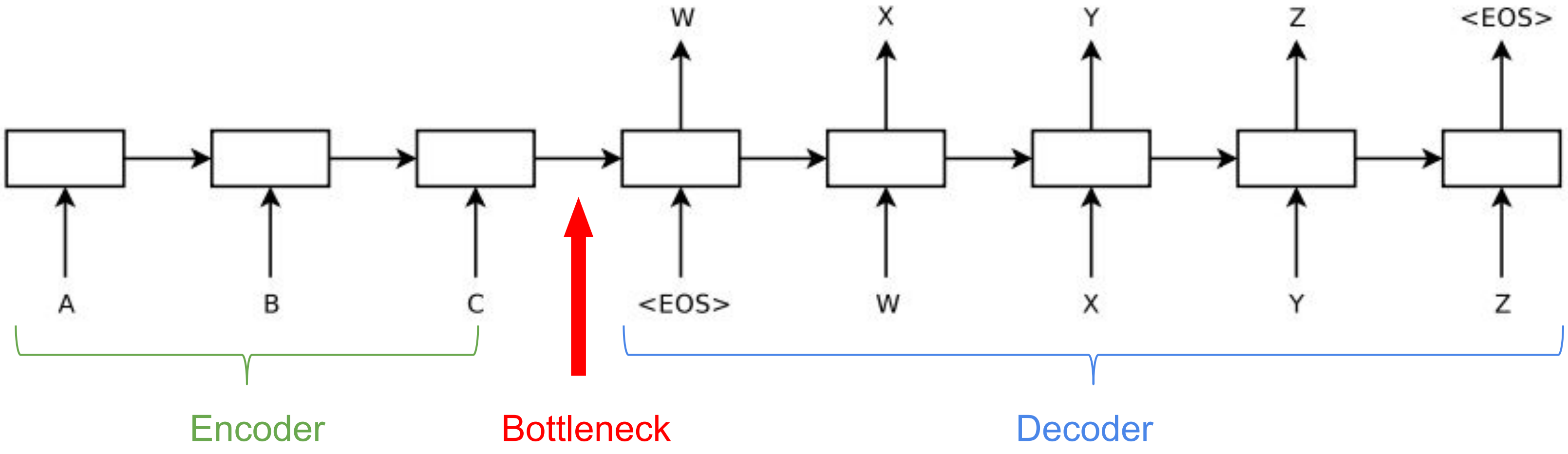
Plan

- ***Weak Supervision NLP***
  - NLP & software 2.0
  - Semantics as a foreign language
  - Semantic search
  - Semantic Parsing Tasks -- WebQuestionsSP, WikiTableQuestions
- ***Neural Symbolic Machines*** (ACL 2017)
  - Compositionality (short term memory)
  - Scalable KB inference (symbolism)
- ***Memory Augmented Policy Optimization*** (NIPS 2018)
  - Experience replay & optimal updating strategy
  - RL vs MML vs ML

[Hochreiter & Schmidhuber 1997]
[Sutskever, Vinyals, Le 2014]

A bit background on Seq2Seq

- Separate a sequence model in to encoder and decoder
- Improves a phrase-based SMT system by re-ranking top candidates
- Cannot perform well by itself due to the information bottleneck



Neural Symbolic Machines

ACL [Liang+ 2017]

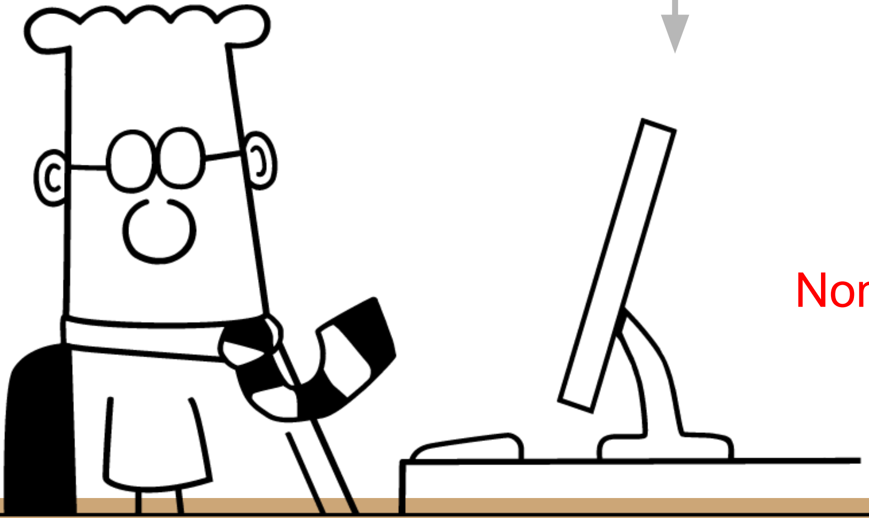
Weak supervision | **Neural** | **Symbolic**

Manager → Question → **Programmer** → Program → **Computer**

Computer → Results → **Programmer** → Answer → **Manager**

Knowledge Base → **Computer**

Predefined Functions → **Computer**



Abstract
Scalable
Precise
Non-differentiable

Computer with Domain Specific Languages

ACL [Liang+ 2017]

- Lisp Interpreter
  - Program => exp_1 exp_2 … exp_n <END>
  - Exp => (f arg_1 arg_2 … arg_n)

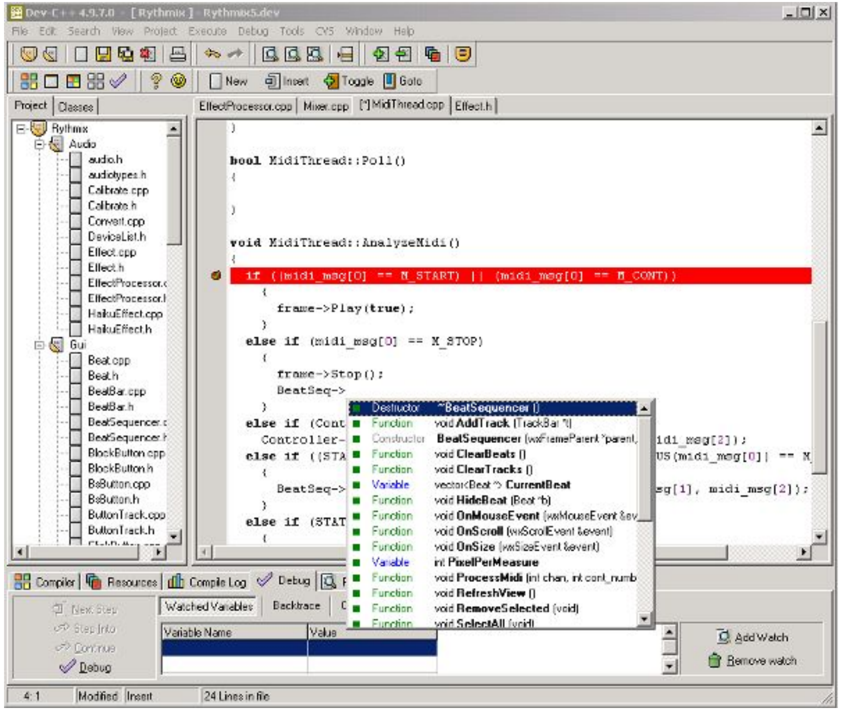
```
(hop m.russell_wilon /education)
(hop v0 /institution)
(filter_= v1 m.univeristy
          /notable_types)
<END>
```

- What functions will be useful for the given task?
  - 10 operations for WebQuesitons
  - 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

Code Assistance to Prune Search Space

ACL [Liang+ 2017]

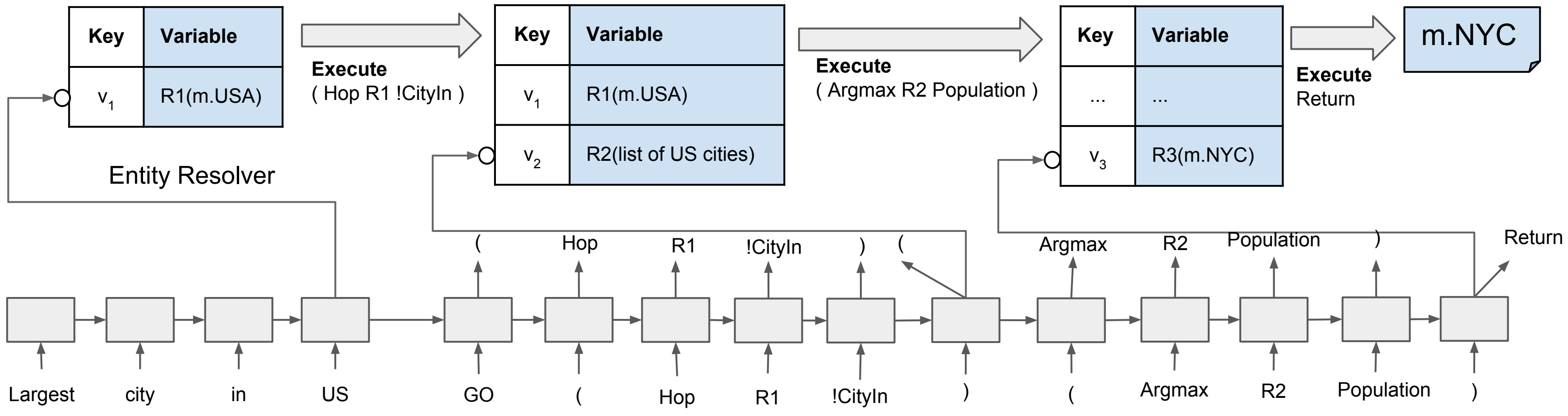
Pen and paper



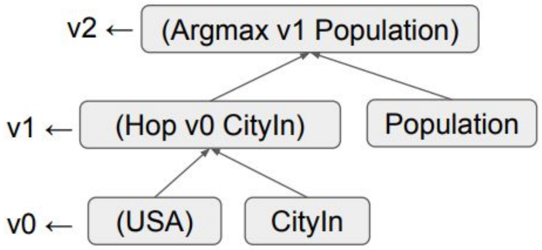
IDE

A lot of computations!

Key-Variable Memory for Semantic Compositionality



- Equivalent to a linearised bottom-up derivation of the recursive program



Augmented REINFORCE

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

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**

State-of-the-Art on WebQuestionsSP

- First end-to-end neural network to achieve SOTA on semantic parsing with weak supervision over large knowledge base
- The performance is approaching 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

Plan

- ***Weak Supervision NLP***
  - NLP & software 2.0
  - Semantics as a foreign language
  - Semantic search
  - Semantic Parsing Tasks -- WebQuestionsSP, WikiTableQuestions

- ***Neural Symbolic Machines*** (ACL 2017)
  - Compositionality (short term memory)
  - Scalable KB inference (symbolism)

- ***Memory Augmented Policy Optimization*** (NIPS 2018)
  - Experience replay & optimal updating strategy
  - RL vs MML vs ML

What is RL?

Directly Optimizing The Expected Reward

- ML optimizes the log likelihood of target sequences

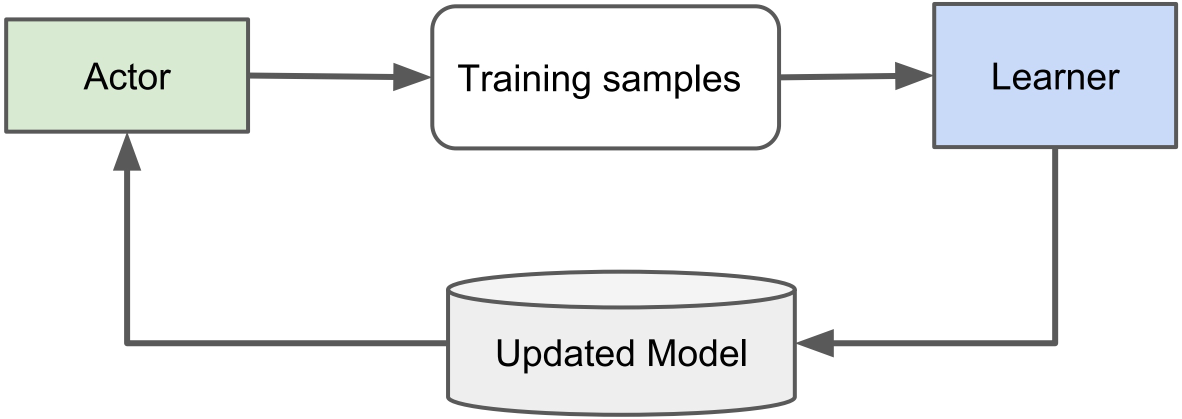
$$J^{ML}(\theta) = \sum_q \log P(a_{0:T}^{best}(q)|q, \theta)$$

- RL optimizes the expected reward under a stochastic policy

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

[Williams 1992]
[Sutton & Barto 1998]

RL models generate their own training data



- Training sample management issue
  - Too many low quality examples ⇒ slow training
  - Boosting high reward experience ⇒ biased training

Challenges of applying RL

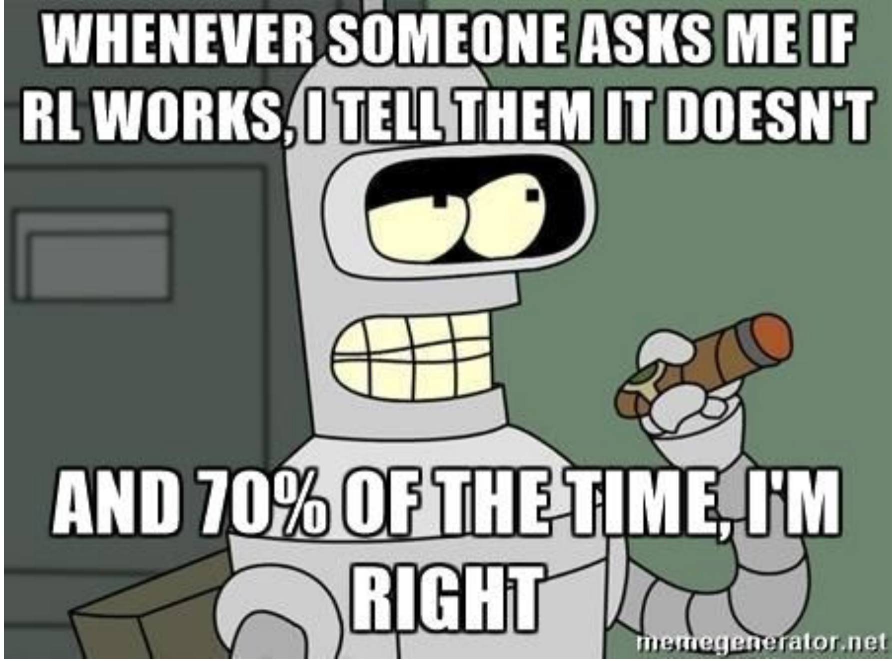
Book [Sutton & Barto 1998]
NIPS [Abbeel & Schulman 2016]

- Large search space (sparse rewards)
  - Supervised pretraining (MLE)
  - Systematic exploration [Houthooft+ 2017]
  - Curiosity [Schmidhuber 1991][Pathak2017]
- Credit assignment (delayed reward)
  - Bootstrapping
    - E.g., AlphaGo uses a value function to estimate the future reward
  - Rollout n-steps
- Train speed & stability (optimization)
  - Trust region approaches (e.g., PPO)
  - **Experience replay** ⇐ **Our focus today**

Efficiency challenge

- RL is still far from data efficient
  - E.g. the best learning algorithm (DeepMind RainbowDQN) “passes median human performance on 57 Atari games at about 18 million frames (around 90 hours) of gameplay, while most humans can pick up a game within a few minutes.”

- How to improve its efficiency?



Credit to Alex Irpan 2018 from Google Brain Robotics

Details

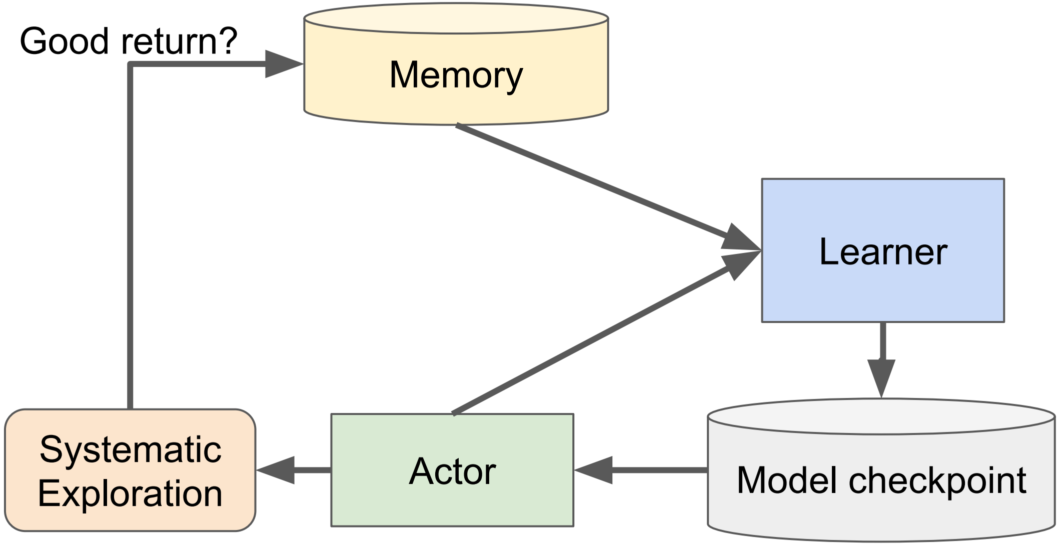
Applying RL to NLP

- Benefits of RL
  - Weak supervision (e.g., expected answer, user clicks)
  - Directly optimizing the metric (e.g., F1, accuracy, BLEU etc.)
  - Work with structured hidden variables (e.g., logical forms/programs)

- Challenges with existing solutions
  - Large search space sparse reward often leads to slow and unstable training
  - Spurious reward often lead to biased solutions

NIPS [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+ 2016]
[Abolafia+ 2017]

Augment REINFORCE with Memory

Linear combination of maximum likelihood objective and expected return:

λ log-likelihood on a top-k buffer + (1 - λ) expected return

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

- Not robust against spurious programs
- The composite objective is ad-hoc, and the gradient is biased

Spurious programs: right answer, wrong reason

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

Which nation won the most silver medal?

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

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

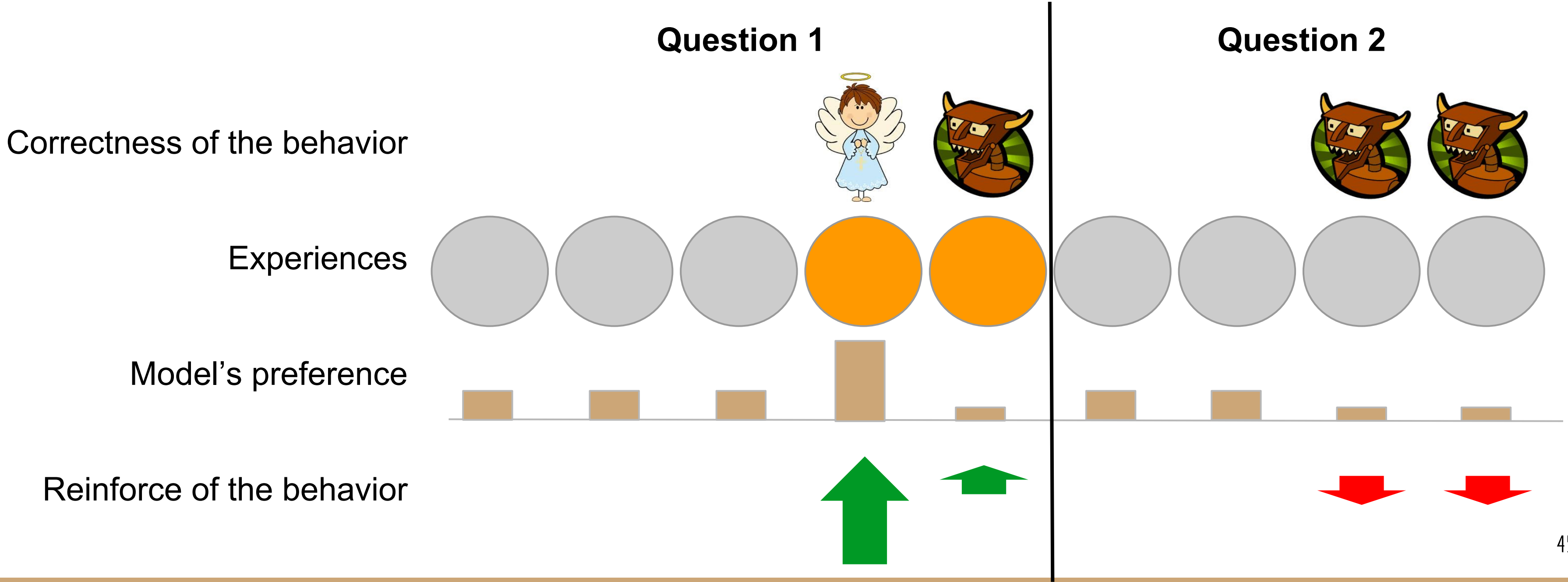
  (argmax rows “Bronze”)
  (hop v1 “Nation”)

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

  ...

Combating spurious rewards

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



Lower the gradient variance without introducing bias

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

re-express expected return as,

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

- **Importance sampling**
  - Sample more frequently inside the buffer
  - Rejection sampling for samples outside the buffer.

Details

NIPS [Liang+ 2018]

Optimal Sample Allocation

Given that we want to apply stratified sampling to estimate the gradient of REINFORCE with baseline under 1/0 rewards. It can be shown that the optimal strategy is to allocate the **same number of samples** to **reward** vs **no reward** experiences

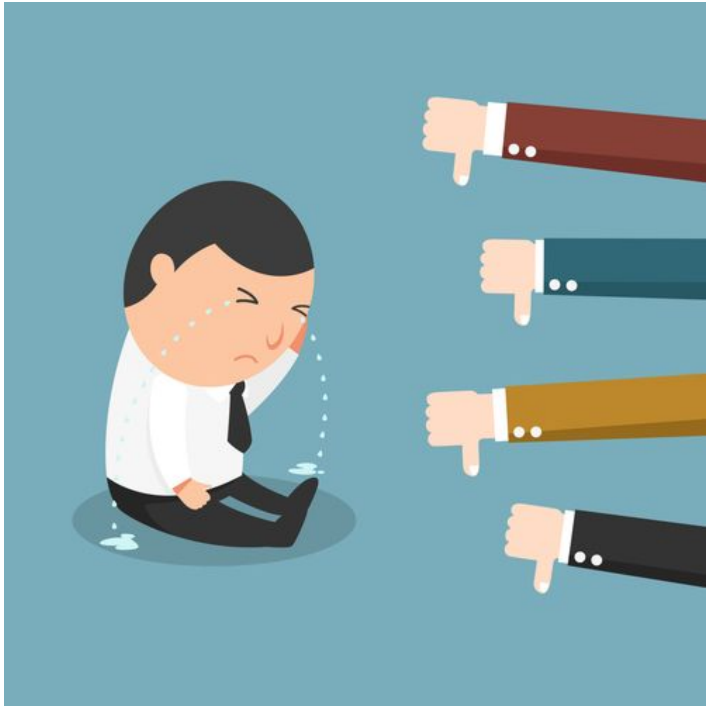
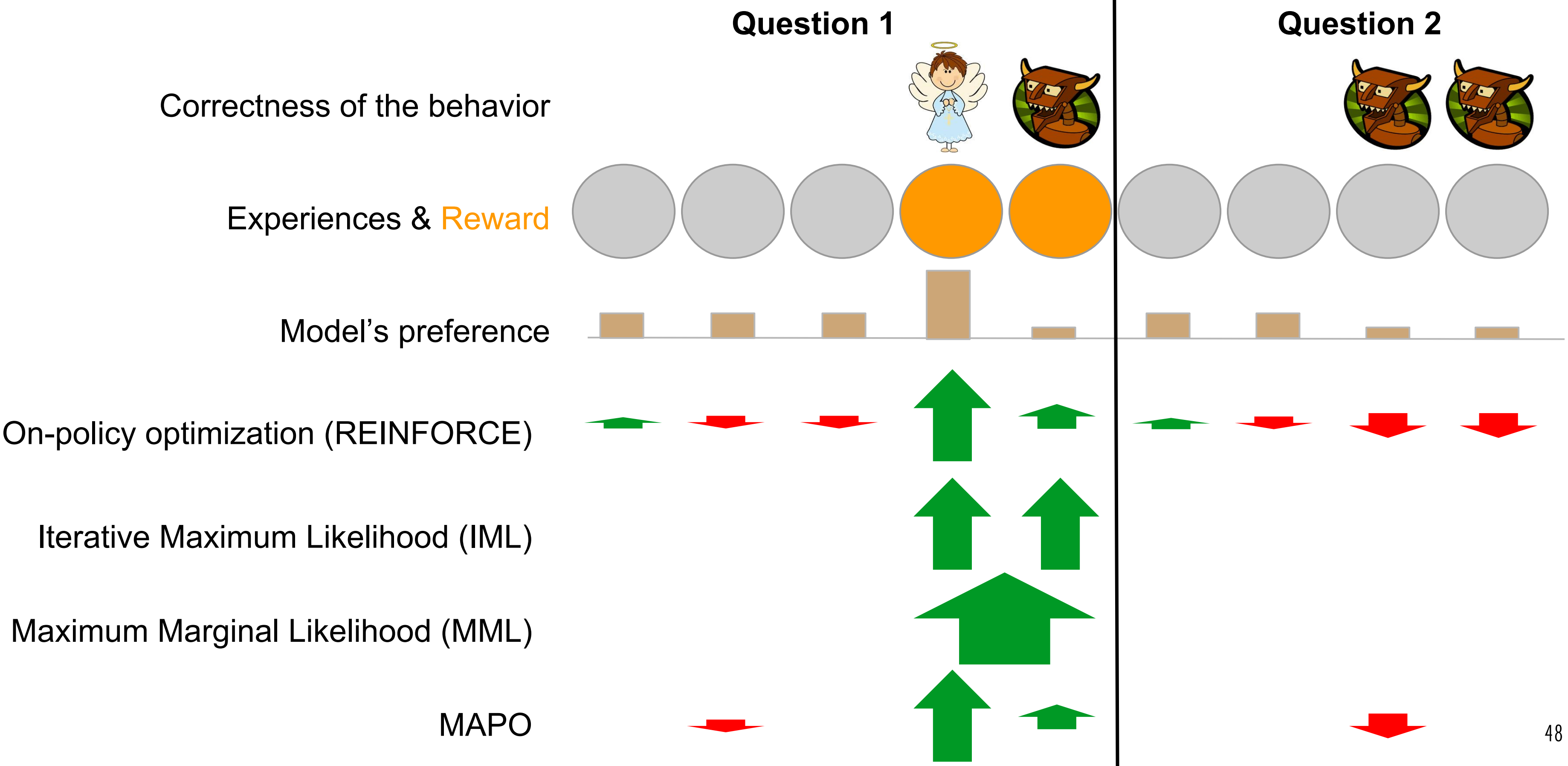


Image source: Guy Harris, 2018
How to Give Feedback in a Non-Threatening Way

Comparison of model update strategies

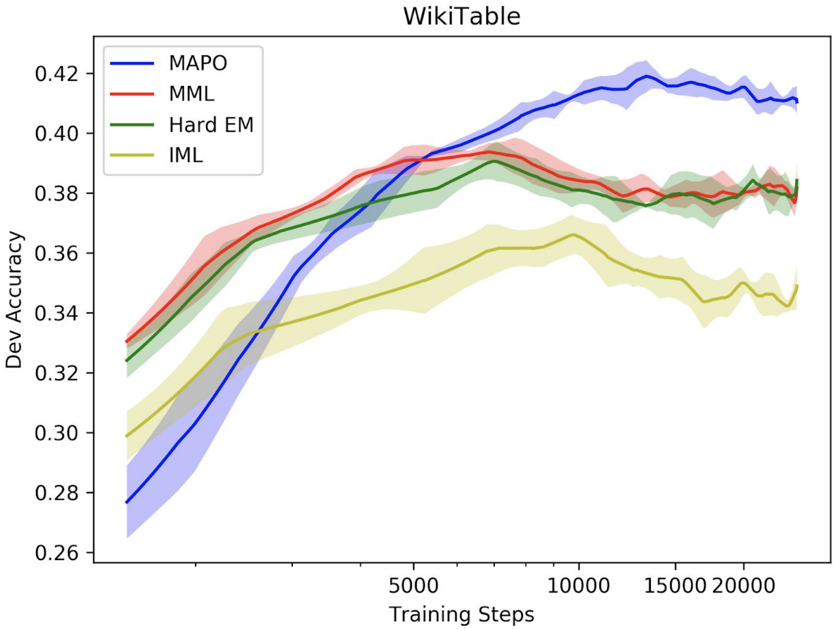
NIPS [Liang+ 2018]

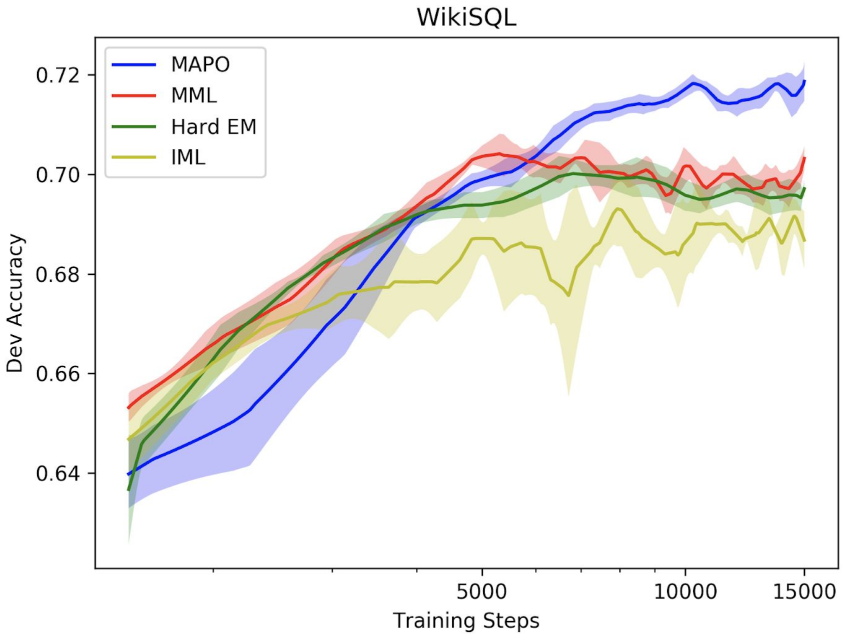


NIPS [Liang+ 2018]

Comparison

- REINFORCE does not work at all
- MAPO is slower but less biased





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

SOTA results with weak supervision

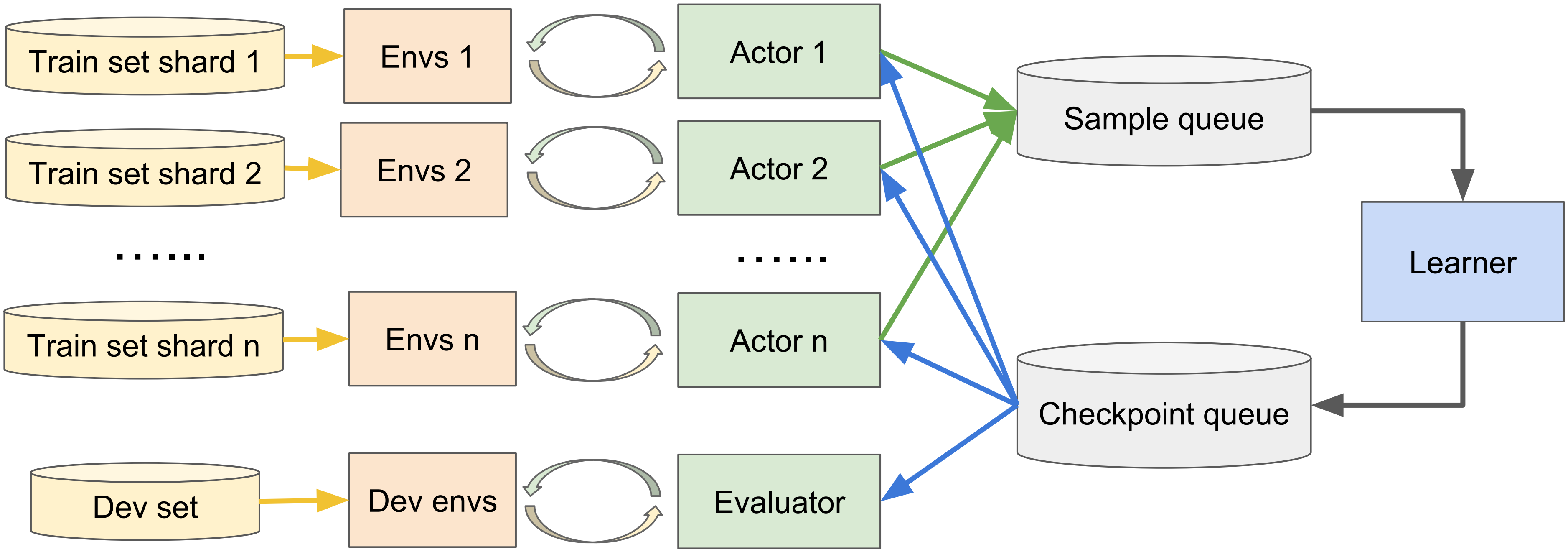
Model	E.S.	Dev.	Test
Pasupat & Liang (2015)[28]	-	37.0	37.1
Neelakantan et al. (2017)[26]	1	34.1	34.2
Neelakantan et al. (2017)[26]	15	37.5	37.7
Haug et al. (2017)[15]	1	-	34.8
Haug et al. (2017)[15]	15	-	38.7
Zhang et al. (2017)[51]	-	40.4	43.7
MAPO	1	42.7	43.8
MAPO (ensembled)	5	-	46.2

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.

Scale up: Distributed Actor-Learner architecture



[Liang et al, 2017; Espeholt et al, 2018; Liang et al, 2018]

Thanks!

- ***Weak Supervision NLP***
  - NLP & software 2.0
  - Semantics as a foreign language
  - Semantic search
  - Semantic Parsing Tasks -- WebQuestionsSP, WikiTableQuestions

- ***Neural Symbolic Machines*** (ACL 2017)
  - Compositionality (short term memory)
  - Scalable KB inference (symbolism)

- ***Memory Augmented Policy Optimization*** (NIPS 2018)
  - Experience replay & optimal updating strategy
  - RL vs MML vs ML