Memory Augmented Policy Optimization for Program Synthesis and Semantic Parsing

- Chen Liang, Mohammad Norouzi, Jonathan Berant, Quoc Le, Ni Lao

Makkunda Sharma

Slides taken from :

- Chen Liang (https://crazydonkey200.github.io/NSM-ACL2017.pdf)
- Panupong Pasupat (https://ppasupat.github.io/resource/ACL2015-slides.pdf)

Images and figures from :

- Memory Augmented Policy Optimization for Program Synthesis and Semantic Parsing

(https://papers.nips.cc/paper/8204-memory-augmented-policy-optimization-for-program-synthesis-and-semantic-parsing.pdf)

Task and Dataset Description

Source : https://ppasupat.github.io/resource/ACL2015-slides.pdf

Task Description

Input: utterance x and HTML table t

Output: answer y

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 = Greece held its last Summer Olympics in which year?

y = 2004

Task Description

Input: utterance x and HTML table t

Output: answer y

Training data: list of (x, t, y) — no logical form

Tables in test data are **not seen** during training

- The model must generalize to unseen table schemas!

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

https://en.wikipedia.org/wiki/Piotr_Kędzia

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia
- Questions x and answers y are from Mechanical Turk — Prompts are given to encourage compositionality

How many …	… ___est …	… last …
… above …	… same … as …	… difference …
… or …	… his …	Requires counting

etc.

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia
- Questions x and answers y are from Mechanical Turk — Prompts are given to encourage compositionality

Prompt: The question must contains "last" (or a synonym)

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

How long did it take this competitor to finish the 4x400 meter relay at Universiade in 2005?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

Where was the competition held immediately before the one in Turkey?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

How many times has this competitor placed 5th or better in competition?

Dataset

WikiTableQuestions dataset:

- 2100 tables
  - Average: 6.3 columns /27.5 rows
- 22000 examples

Baseline Description

Source : https://crazydonkey200.github.io/NSM-ACL2017.pdf

Neural Symbolic Machines

Semantic Parsing on Freebase with Weak Supervision

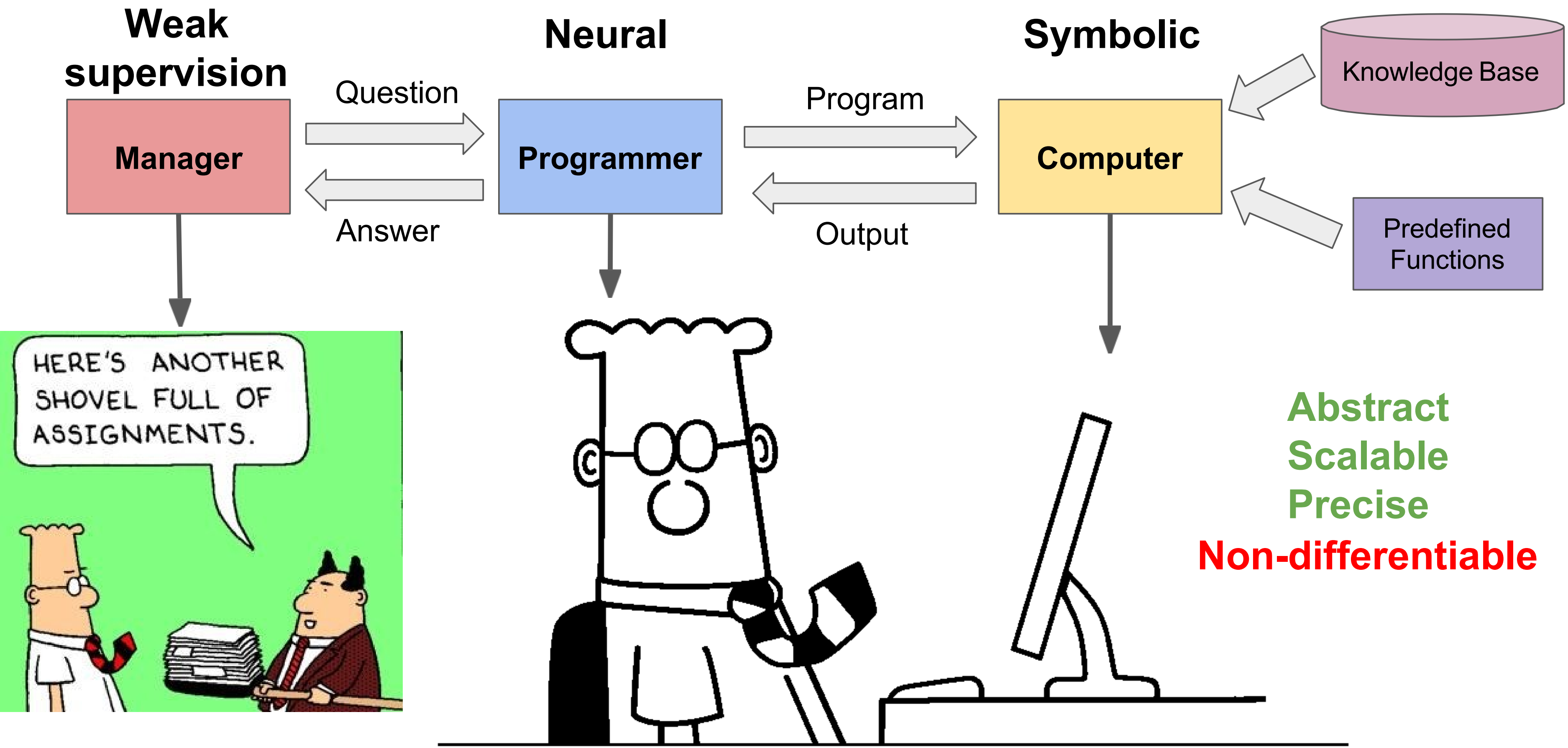
Chen Liang, Jonathan Berant, Quoc Le, Kenneth Forbus, Ni Lao

Research at Google

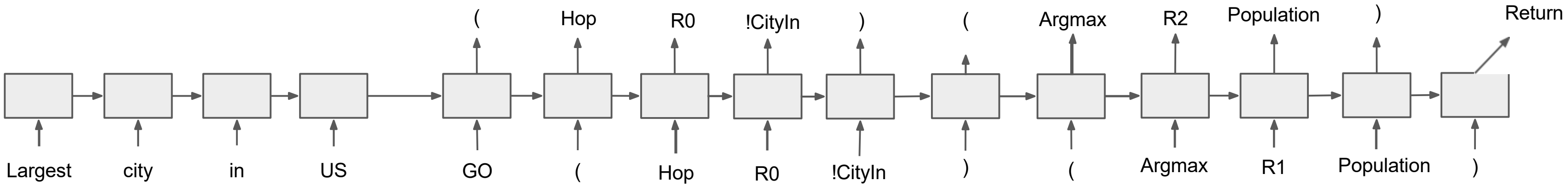
NORTHWESTERN UNIVERSITY

TEL AVIV UNIVERSITY

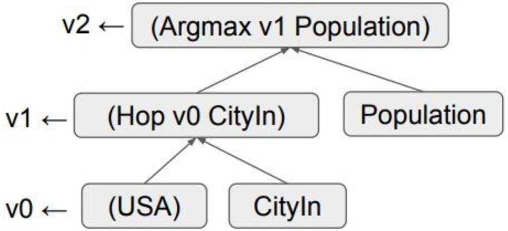
Neural Symbolic Machines



Simple Seq2Seq model is not enough



1. Compositionality



2. Large Search Space

23K predicates,
82M entities,
417M triplets

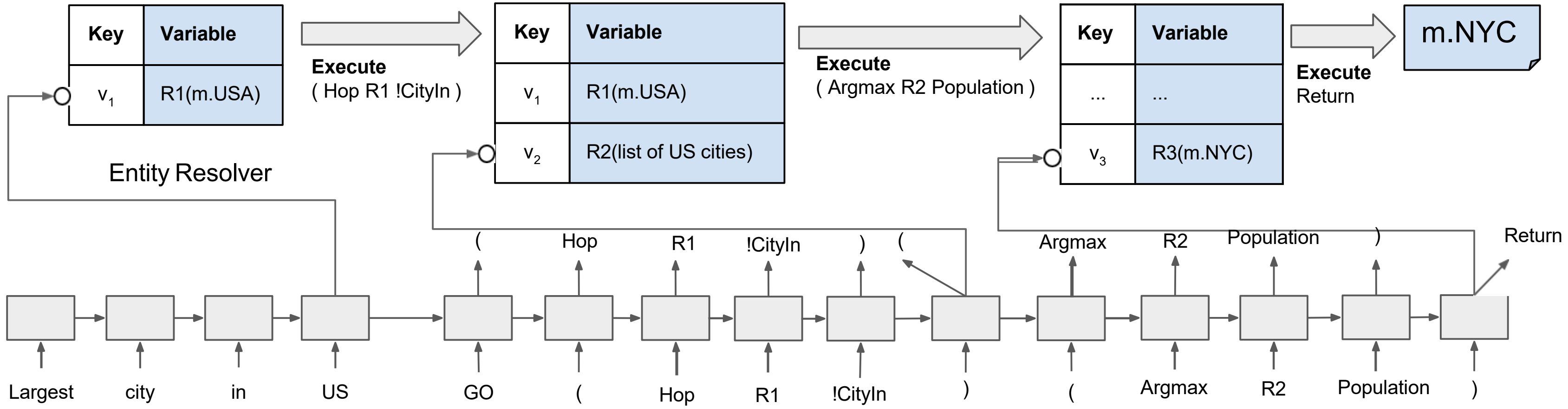
1.Key-Variable Memory

2.Code Assistance
3.Augmented REINFORCE

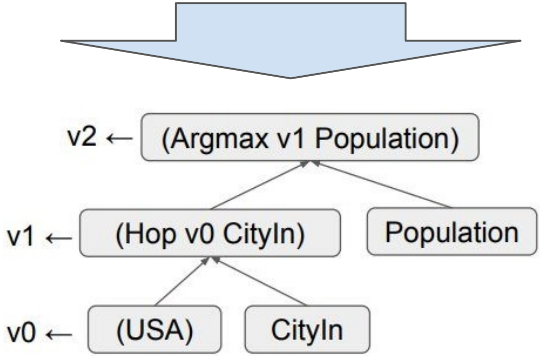
Overview

- Motivation: Semantic Parsing and Program Induction
- Neural Symbolic Machines
  - Key-Variable Memory
  - Code Assistance
  - Augmented REINFORCE
- Experiments and analysis

Key-Variable Memory for Compositionality

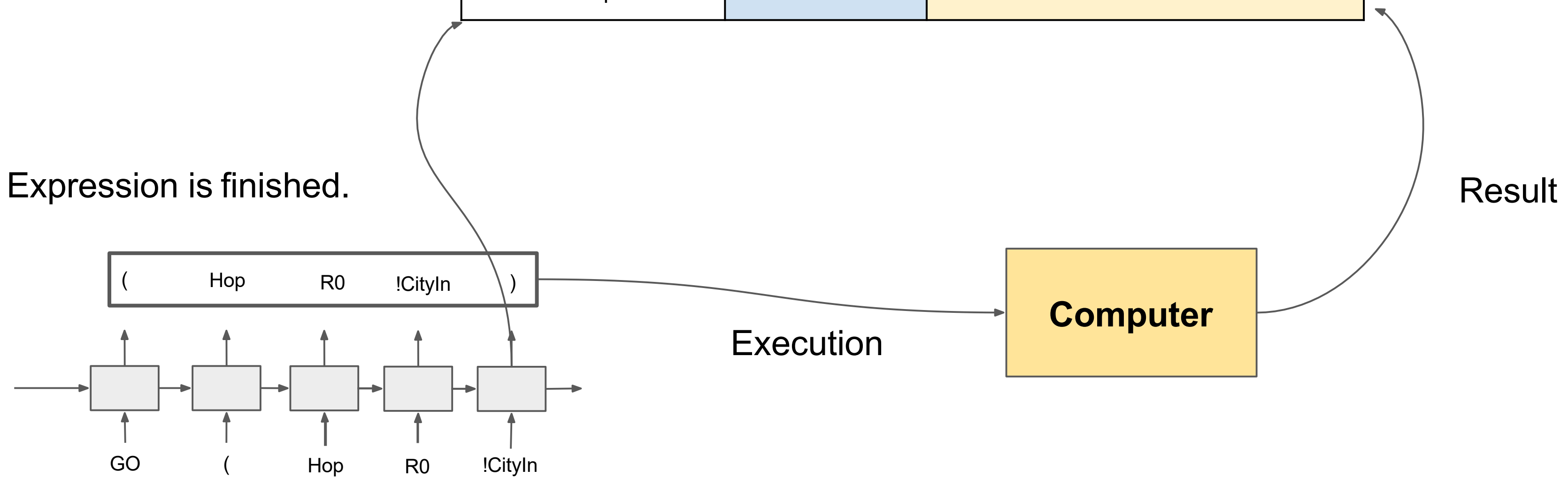


- A linearised bottom-up derivation of the recursive program.

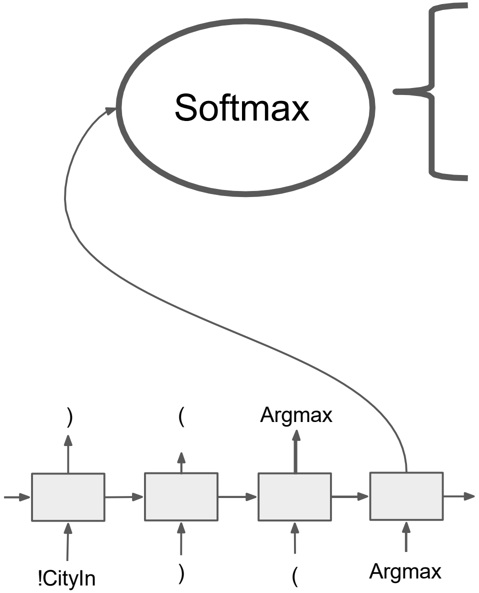


Key-Variable Memory: Save Intermediate Value

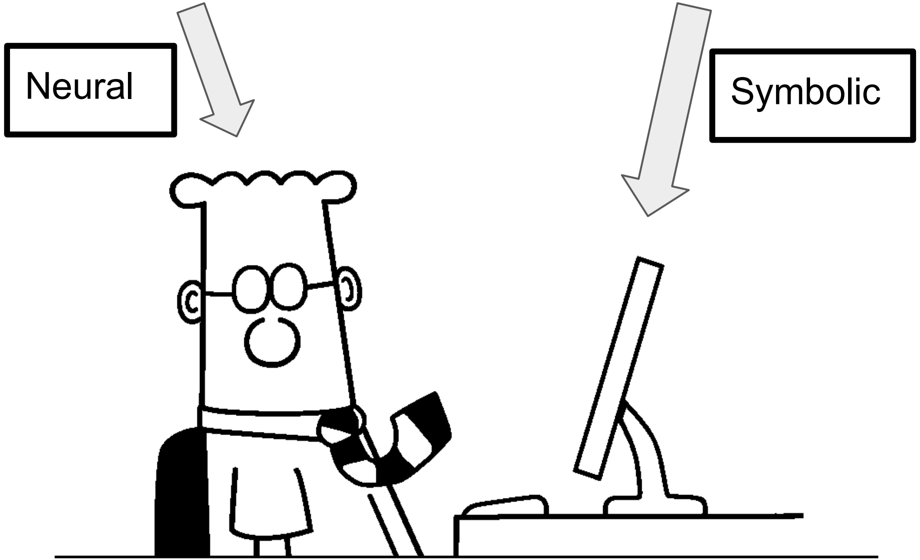
Key (Embedding)	**Variable** (Symbol)	**Value** (Data in Computer)
V_0	R0	m.USA
V_1	R1	[m.SF, m.NYC, ...]



Key-Variable Memory: Reuse Intermediate Value



Key (Embedding)	**Variable** (Symbol)	**Value** (Data in Computer)
V_0	`R0`	`m.USA`
V_1	`R1`	`[m.SF, m.NYC, ...]`

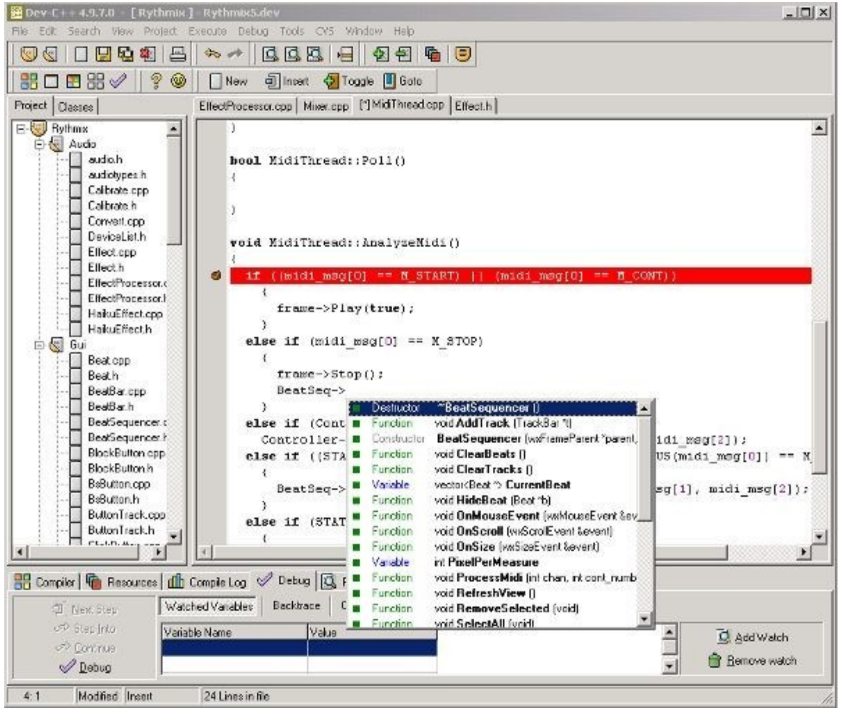


Overview

- Motivation: Semantic Parsing and Program Induction
- Neural Symbolic Machines
  - Key-Variable Memory
  - Code Assistance
  - Augmented REINFORCE
- Experiments and analysis

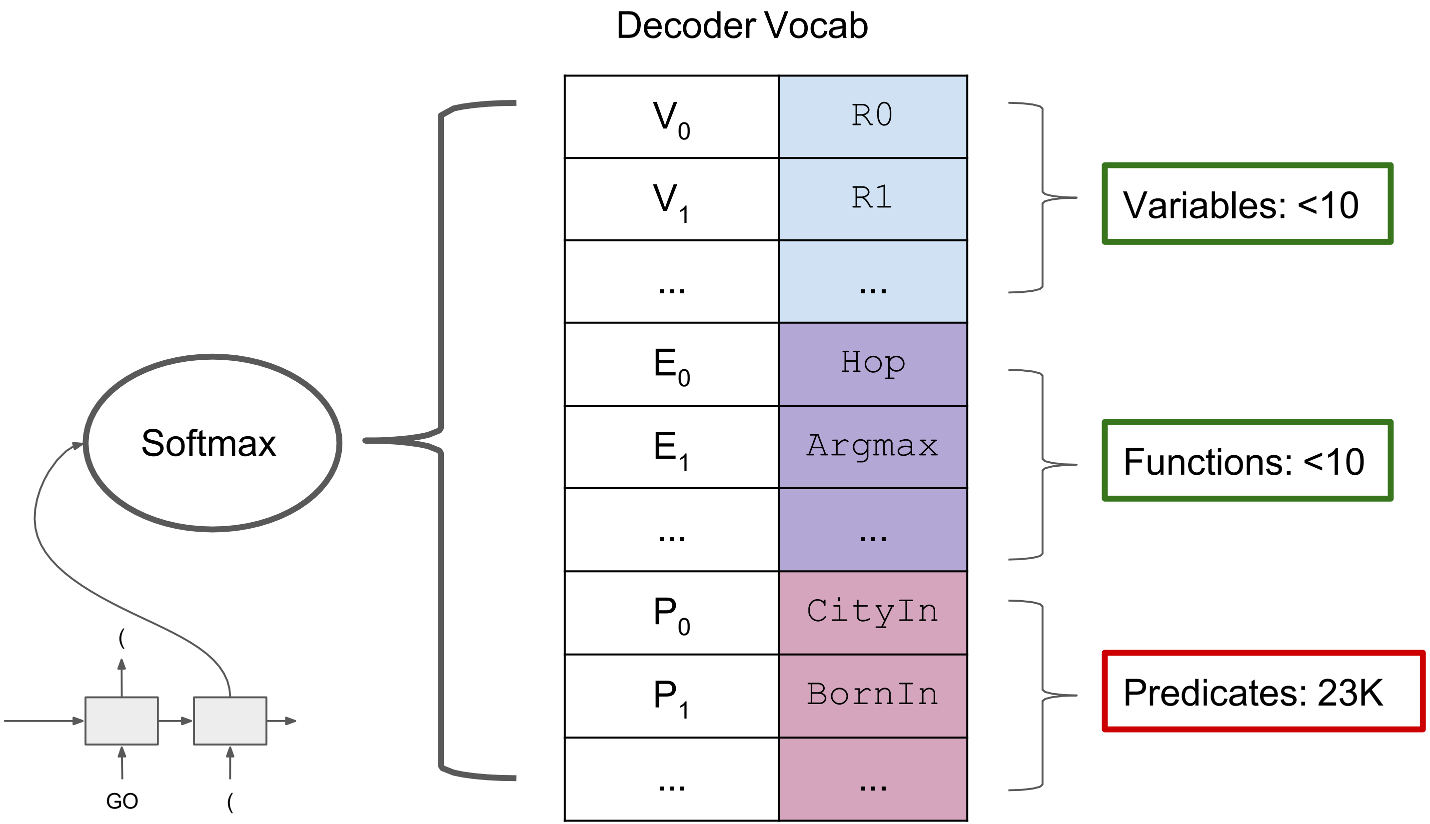
Code Assistance: Prune Search Space

Pen and paper



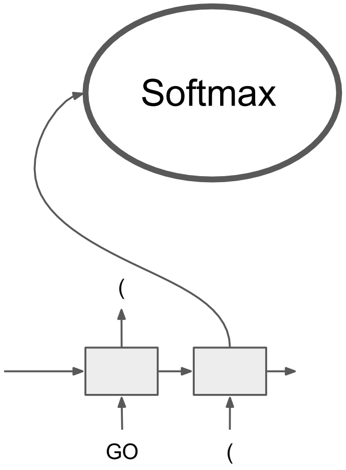
IDE

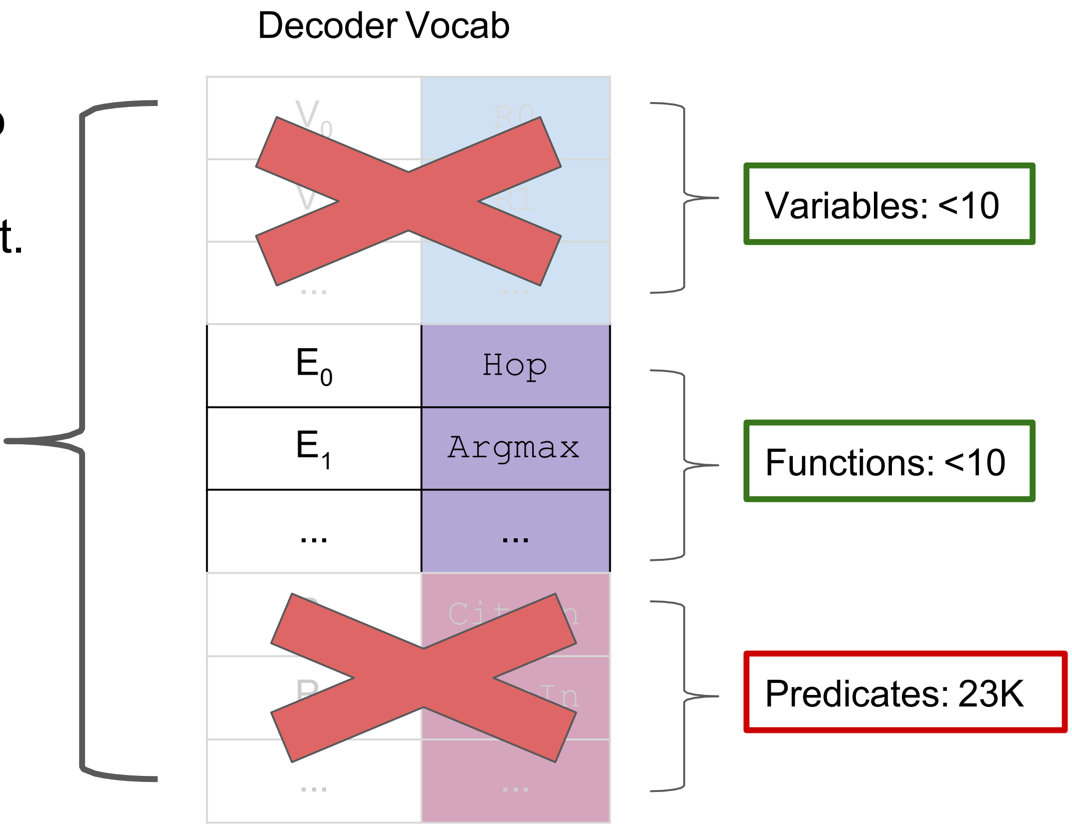
Code Assistance: Syntactic Constraint



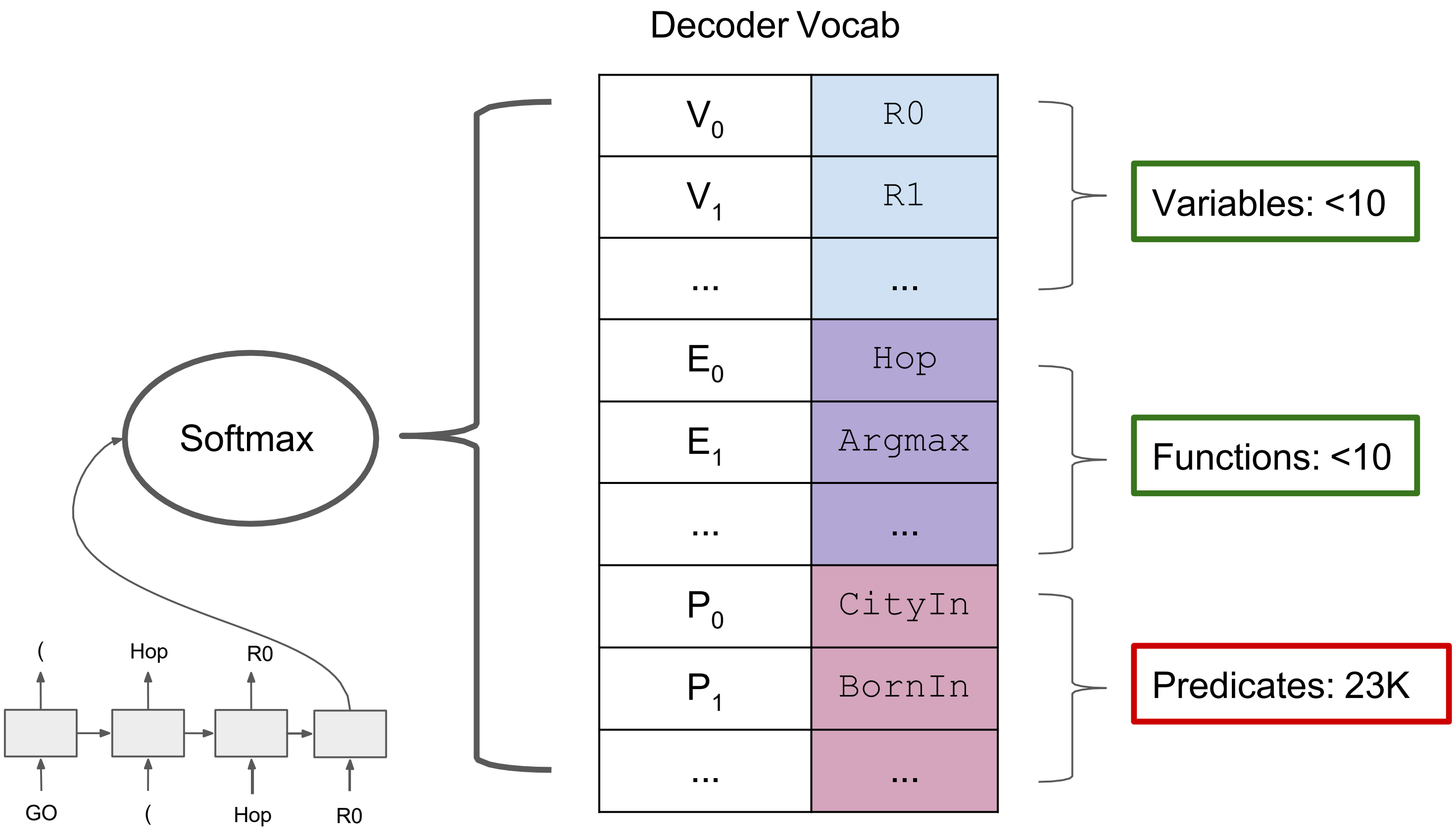
Code Assistance: Syntactic Constraint

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



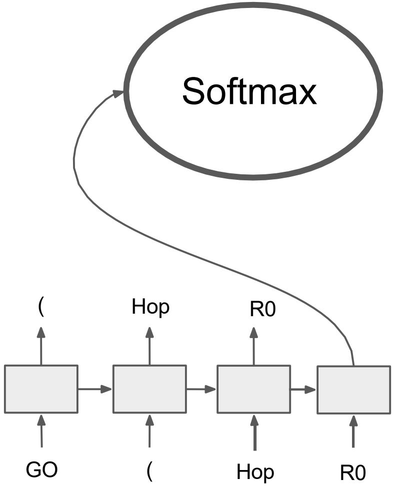


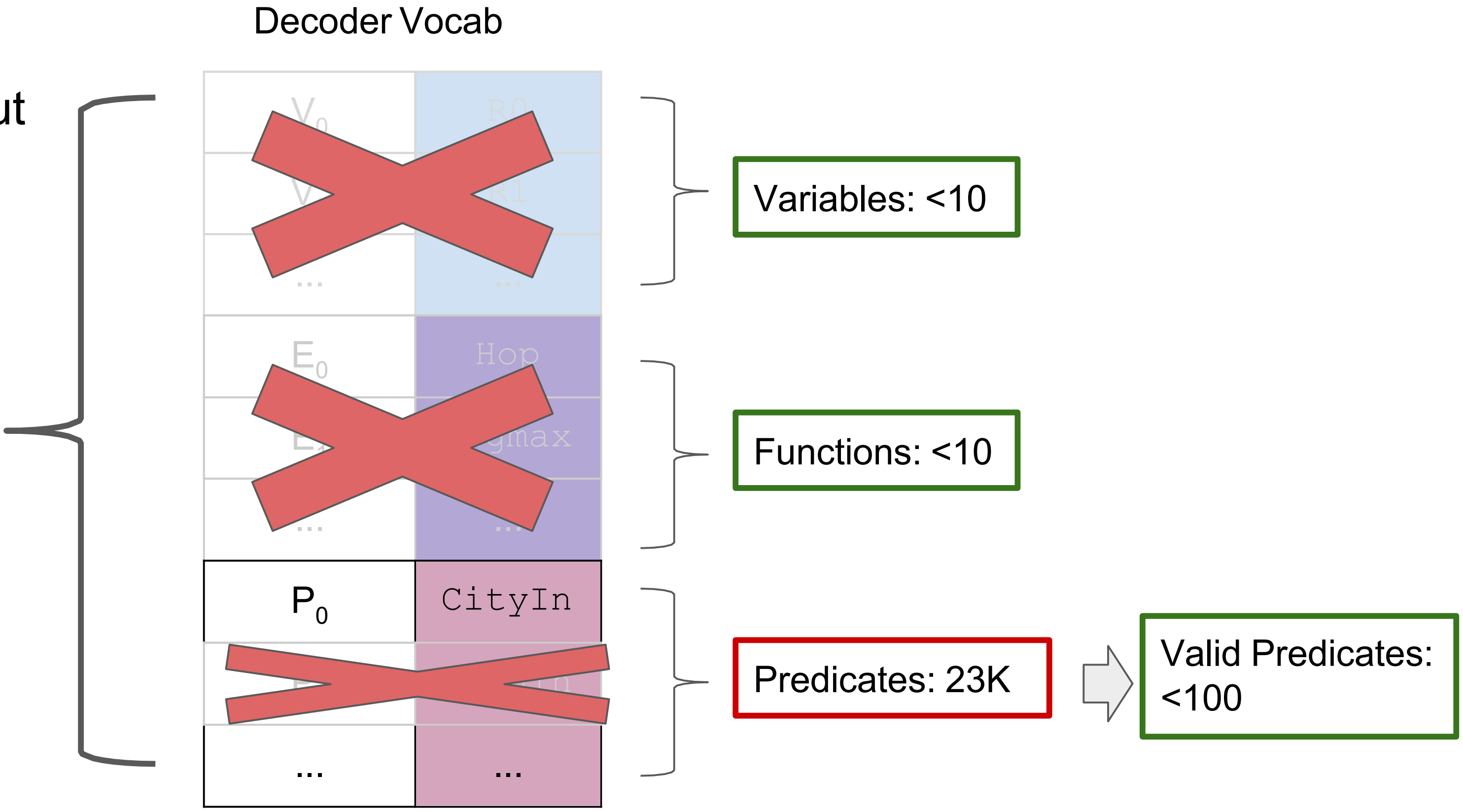
Code Assistance: Semantic Constraint



Code Assistance: Semantic Constraint

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

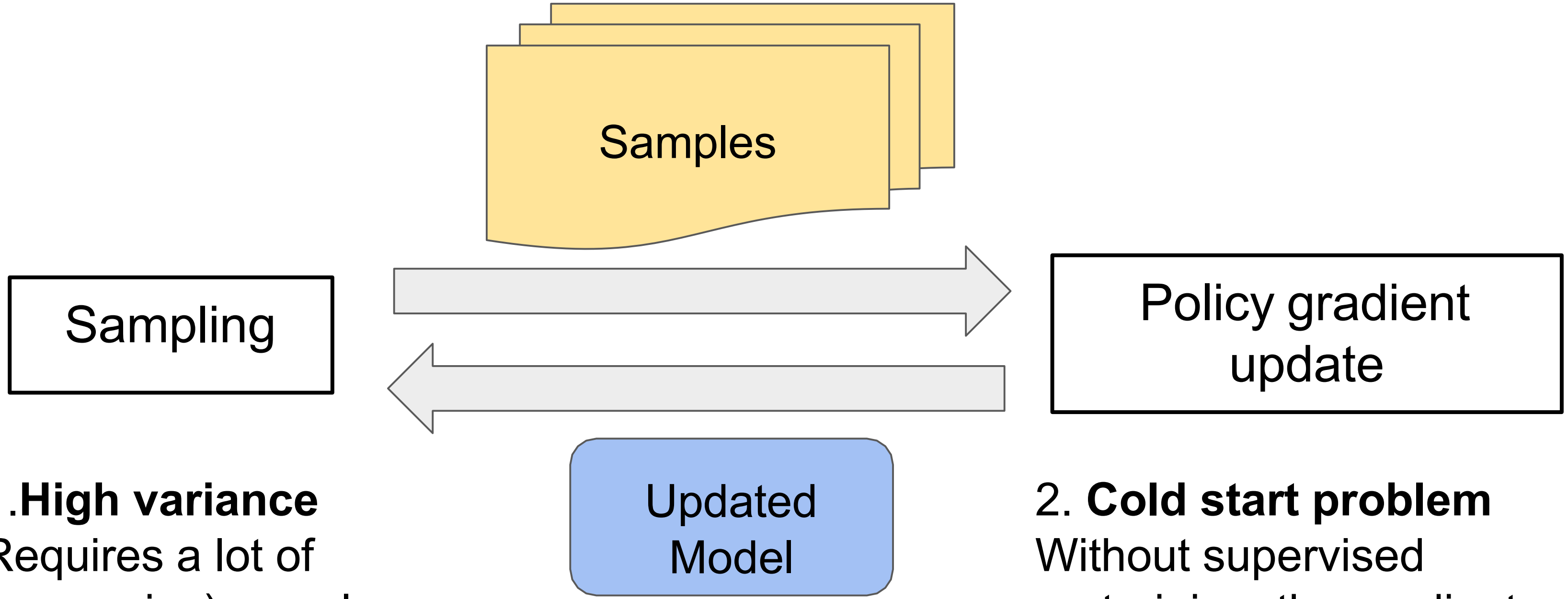




Overview

- Motivation: Semantic Parsing and Program Induction
- Neural Symbolic Machines
  - Key-Variable Memory
  - Code Assistance
  - Augmented REINFORCE
- Experiments and analysis

REINFORCE Training

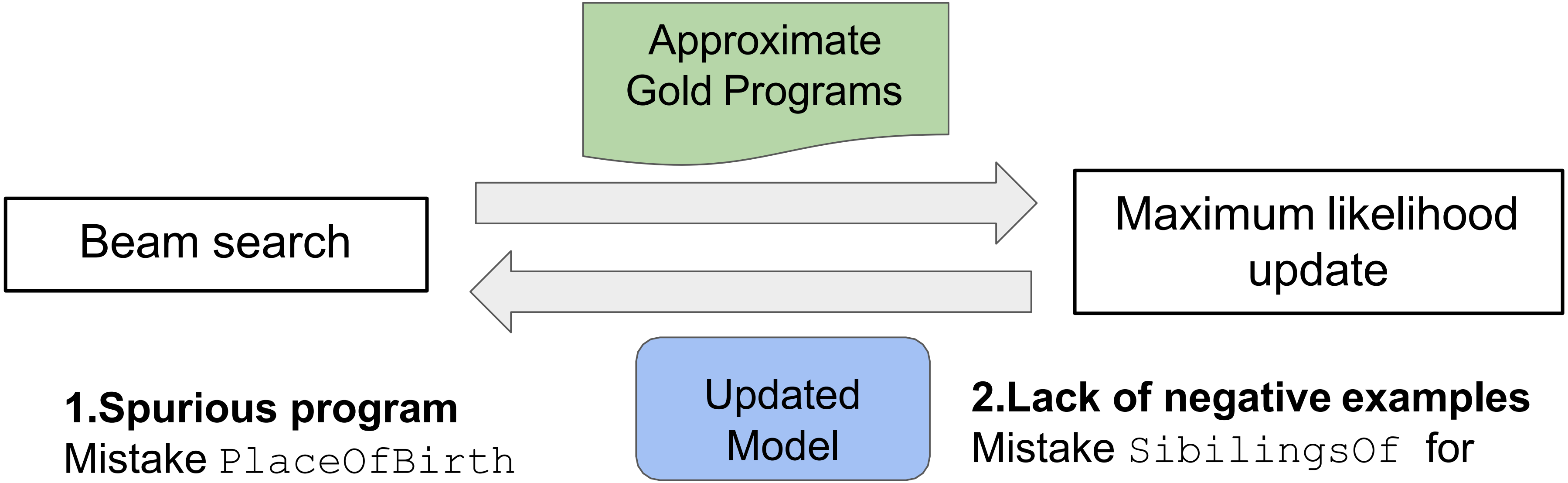


1.**High variance**
Requires a lot of (expensive) samples

2. **Cold start problem**
Without supervised pretraining, the gradients at the beginning

$$\nabla_\theta J^{RL}(\theta) = \sum_q \sum_{a_{0:T}} \boxed{P(a_{0:T}|q,\theta)} [R(q, a_{0:T}) - B(q)] \nabla_\theta \log P(a_{0:T}|q,\theta)$$

Iterative Maximum Likelihood Training (Hard EM)

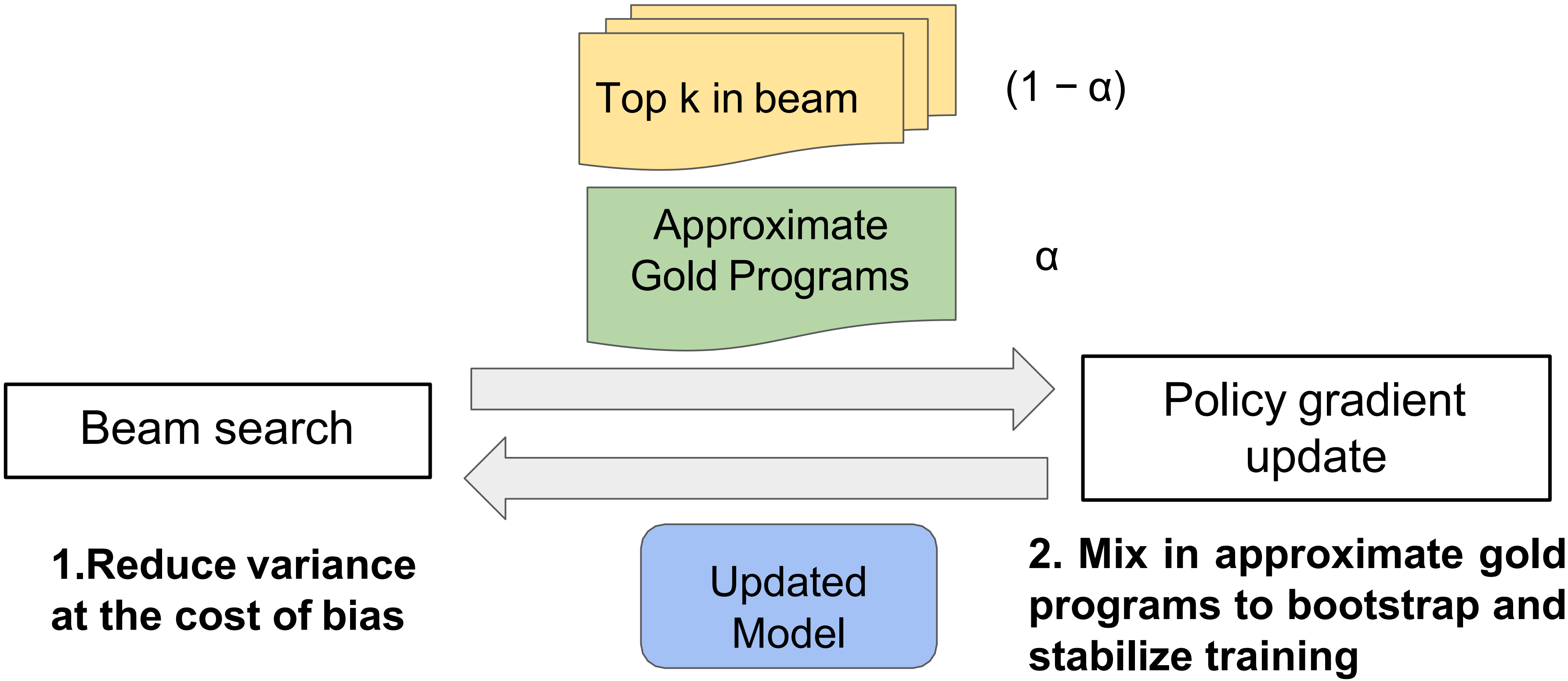


1.Spurious program
Mistake `PlaceOfBirth` for `PlaceOfDeath`.

2.Lack of negative examples
Mistake `SibilingsOf` for `ParentsOf`.

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

Augmented REINFORCE



1.Reduce variance at the cost of bias

2. Mix in approximate gold programs to bootstrap and stabilize training

Paper Description

RL Formulation of problem :

- To learn a mapping from x -> a where x is the question and a is the program generated with weak supervision given in form of correct answer y
- Our reward R(a|x,y) is 1 if answer is correct and 0 otherwise
- We have a distribution pi-theta(a|x) over countable set of all problems denoted by A
- To synthesize a program for a novel context , we find the most likely program under distribution pi-theta by taking argmax (pi-theta(a|x)) over all a in A
- Our objective function is :

$$\mathcal{O}_{\mathrm{ER}}(\theta) = \sum_{\mathbf{a} \in \mathcal{A}} \pi_\theta(\mathbf{a}) \, R(\mathbf{a}) = \mathbb{E}_{\mathbf{a} \sim \pi_\theta(\mathbf{a})} \, R(\mathbf{a}) \,.$$

RL Formulation of problem - cont :

- To estimate the gradient of the expected return we can use REINFORCE using Monte Carlo (MC) samples.
- Using K trajectories sampled from the current policy , the gradient can be estimated as :

$$\nabla_\theta \mathcal{O}_{\mathrm{ER}}(\theta) = \mathbb{E}_{\mathbf{a}\sim\pi_\theta(\mathbf{a})} \nabla \log \pi_\theta(\mathbf{a})\, R(\mathbf{a}) \;\approx\; \frac{1}{K}\sum_{k=1}^{K} \nabla \log \pi_\theta(\mathbf{a}^{(k)})\, [R(\mathbf{a}^{(k)}) - b]$$

- A baseline b is subtracted to reduce variance of gradient estimates

MAPO intuition :

- To reduce variance in gradient estimation and prevent forgetting high reward trajectories people introduce a memory buffer which saves a set of promising trajectories denoted B
- Here the training objective is not directly optimizing the expected return any more because the second term introduces a bias - eg: spurious programs with the correct answer
- MAPO tries to do this in a principled way - by re expressing the objective as a weighted sum of two terms

$$\mathcal{O}_{\mathrm{AUG}}(\theta) = \lambda \mathcal{O}_{\mathrm{ER}}(\theta) + (1 - \lambda) \sum_{\mathbf{a} \in \mathcal{B}} \log \pi_\theta(\mathbf{a})$$

MAPO :

$$\begin{aligned}\mathcal{O}_{\mathrm{ER}}(\theta) &= \sum_{\mathbf{a}\in\mathcal{B}} \pi_\theta(\mathbf{a})\, R(\mathbf{a}) &+& \sum_{\mathbf{a}\in(\mathcal{A}-\mathcal{B})} \pi_\theta(\mathbf{a})\, R(\mathbf{a}) \\ &= \pi_{\mathcal{B}}\, \underbrace{\mathbb{E}_{\mathbf{a}\sim\pi_\theta^+(\mathbf{a})}\, R(\mathbf{a})}_{\text{Expectation inside } \mathcal{B}} &+& (1-\pi_{\mathcal{B}})\, \underbrace{\mathbb{E}_{\mathbf{a}\sim\pi_\theta^-(\mathbf{a})}\, R(\mathbf{a})}_{\text{Expectation outside } \mathcal{B}}\end{aligned}$$

$$\pi_\theta^+(\mathbf{a}) = \begin{cases}\pi_\theta(\mathbf{a})/\pi_{\mathcal{B}} & \text{if } \mathbf{a}\in\mathcal{B} \\ 0 & \text{if } \mathbf{a}\notin\mathcal{B}\end{cases}, \qquad \pi_\theta^-(\mathbf{a}) = \begin{cases}0 & \text{if } \mathbf{a}\in\mathcal{B} \\ \pi_\theta(\mathbf{a})/(1-\pi_{\mathcal{B}}) & \text{if } \mathbf{a}\notin\mathcal{B}\end{cases}$$

$$\nabla_\theta \mathcal{O}_{\mathrm{ER}}(\theta) = \pi_{\mathcal{B}}\, \mathbb{E}_{\mathbf{a}\sim\pi_\theta^+(\mathbf{a})}\, \nabla \log \pi_\theta(\mathbf{a}) R(\mathbf{a}) + (1-\pi_{\mathcal{B}})\, \mathbb{E}_{\mathbf{a}\sim\pi_\theta^-(\mathbf{a})}\, \nabla \log \pi_\theta(\mathbf{a}) R(\mathbf{a})$$

Techniques to improve efficiency of MAPO :

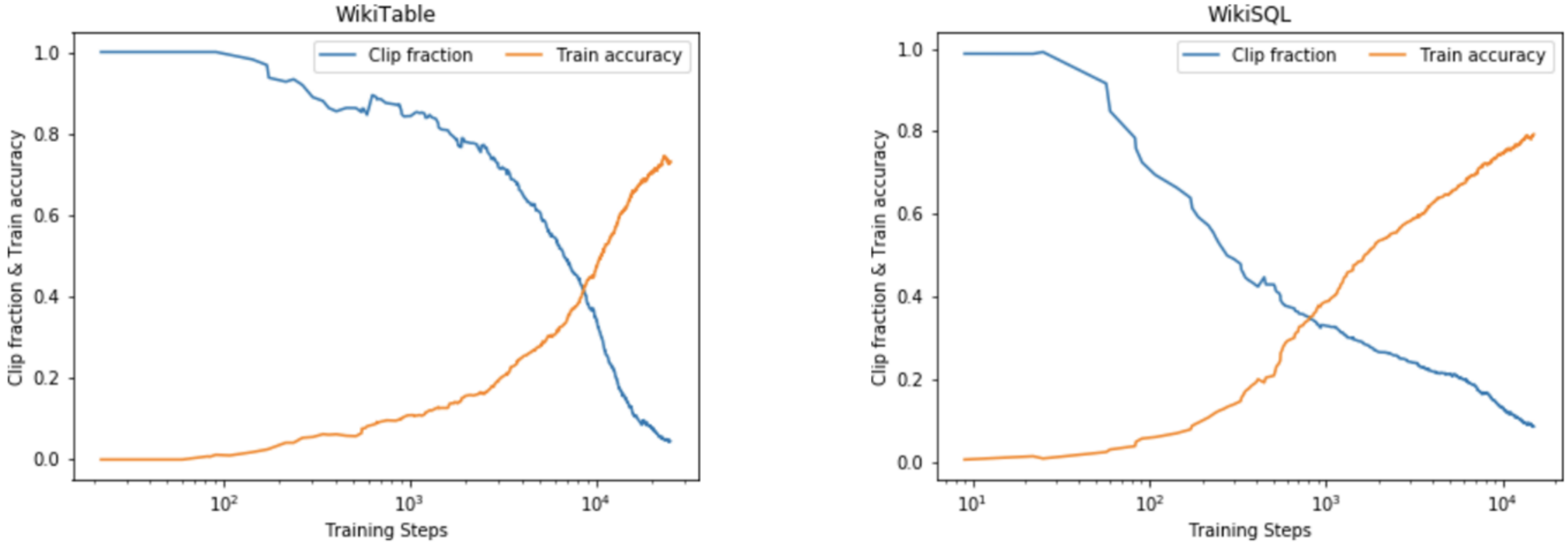
- Memory Weight Clipping
- Systematic Exploration
- Distributed Sampling

Memory Weight Clipping :

- For the cold start problem
- Initially forces the estimates to pay more attention to the high reward trajectories

$$\nabla_\theta \mathcal{O}_{\text{ER}}^c(\theta) = \pi_{\mathcal{B}}^c \, \mathbb{E}_{\mathbf{a} \sim \pi_\theta^+(\mathbf{a})} \nabla \log \pi_\theta(\mathbf{a}) R(\mathbf{a}) + (1 - \pi_{\mathcal{B}}^c) \, \mathbb{E}_{\mathbf{a} \sim \pi_\theta^-(\mathbf{a})} \nabla \log \pi_\theta(\mathbf{a}) R(\mathbf{a})$$

$$\pi_{\mathcal{B}}^c = \max(\pi_{\mathcal{B}}, \alpha)$$

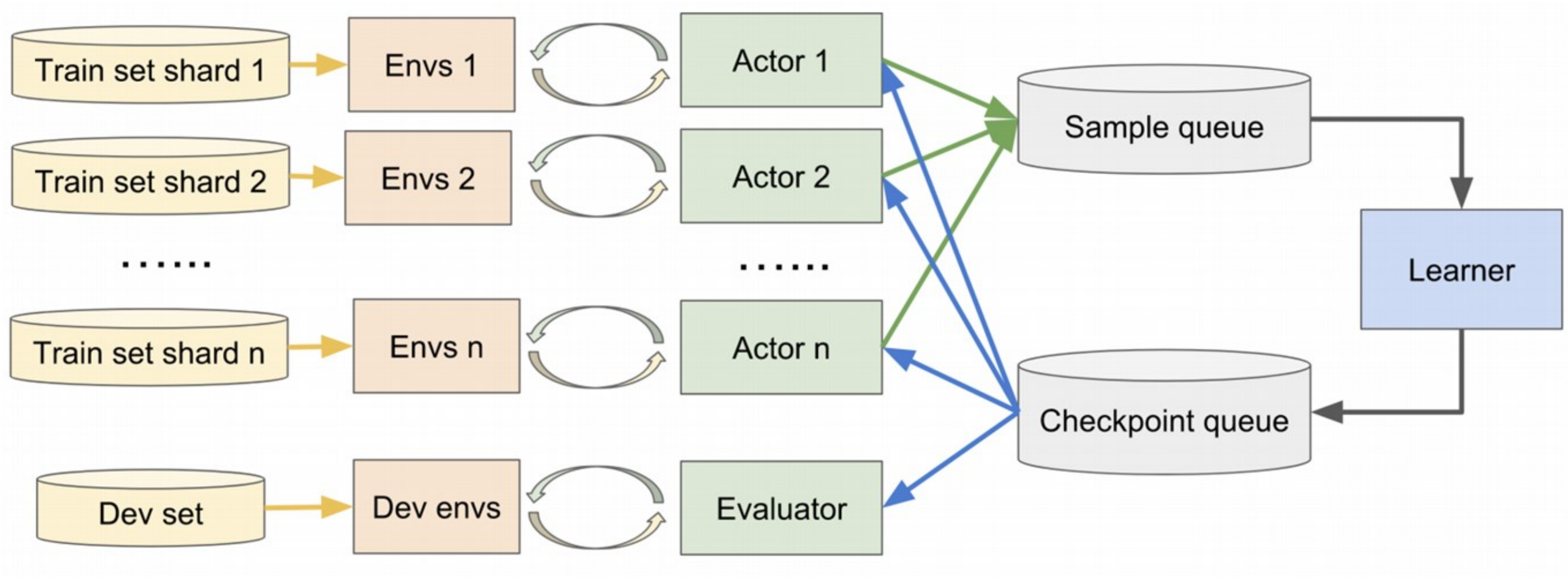

WikiTable
Clip fraction
Train accuracy
Clip fraction & Train accuracy
Training Steps
WikiSQL
Clip fraction
Train accuracy
Clip fraction & Train accuracy
Training Steps

Systematic Exploration :

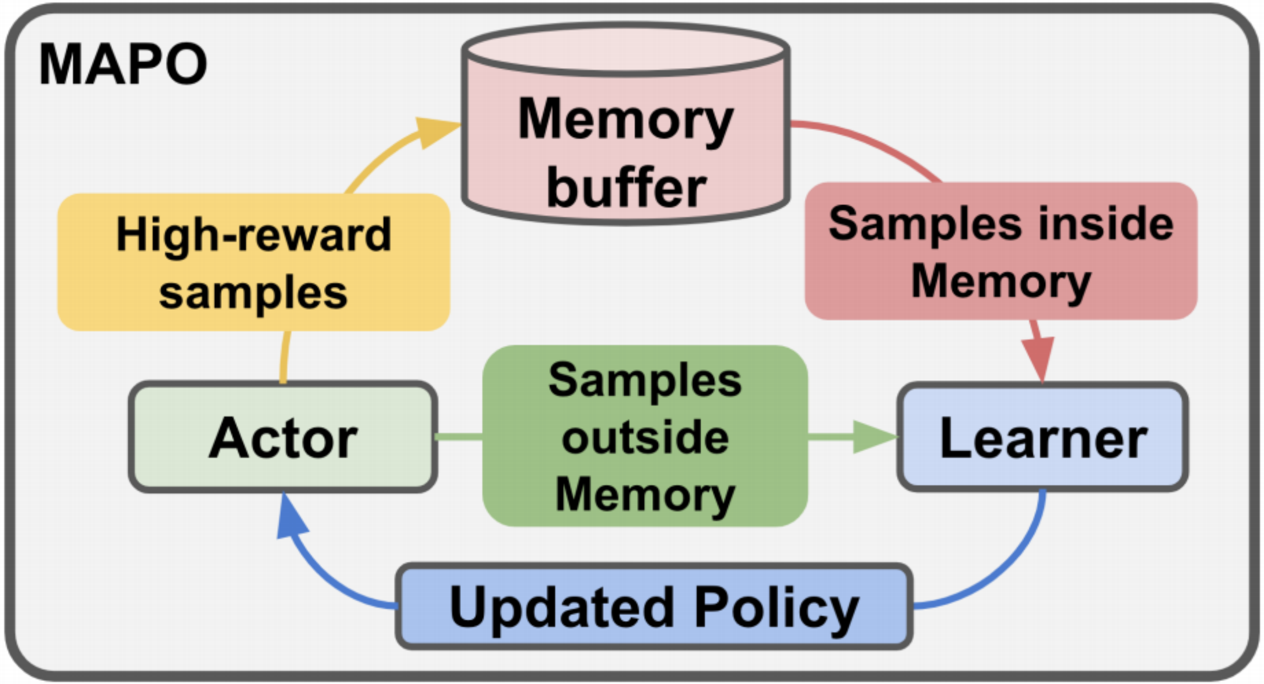
Algorithm 1 Systematic Exploration

Input: context $\mathbf{x}$, policy π, fully explored sub-sequences $\mathcal{B}^e$, high-reward sequences $\mathcal{B}$
Initialize: empty sequence $a_{0:0}$
while true **do**
  $V = \{a \mid a_{0:t-1} \| a \notin B^e\}$
  if $V == \emptyset$ **then**
    $\mathcal{B}^e \leftarrow \mathcal{B}^e \cup a_{0:t-1}$
    break
  sample $a_t \sim \pi^V(a|a_{0:t-1})$
  $a_{0:t} \leftarrow a_{0:t-1} \| a_t$
  if $a_t ==$ EOS **then**
    if $R(a_{0:t}) > 0$ **then**
      $\mathcal{B} \leftarrow \mathcal{B} \cup a_{0:t}$
    $\mathcal{B}^e \leftarrow \mathcal{B}^e \cup a_{0:t}$
    break

Distributed sampling :



Final MAPO :


MAPO
Memory buffer
High-reward samples
Samples inside Memory
Actor
Samples outside Memory
Learner
Updated Policy

MAPO algorithm :

Algorithm 2 MAPO

Input: data $\{(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^N$, memories $\{(\mathcal{B}_i, \mathcal{B}_i^e)\}_{i=1}^N$, constants α, ϵ, M

repeat ▷ **for all actors**

  Initialize training batch $D \leftarrow \varnothing$

  Get a batch of inputs C

  for $(\mathbf{x}_i, \mathbf{y}_i, \mathcal{B}_i^e, \mathcal{B}_i) \in C$ **do**

    $\text{Algorithm1}(\mathbf{x}_i, \pi_\theta^{old}, \mathcal{B}_i^e, \mathcal{B}_i)$

    Sample $\mathbf{a}_i^+ \sim \pi_\theta^{old}$ over $\mathcal{B}_i$

    $w_i^+ \leftarrow \max(\pi_\theta^{old}(\mathcal{B}_i), \alpha)$

    $D \leftarrow D \cup (\mathbf{a}_i^+, R(\mathbf{a}_i^+), w_i^+)$

    Sample $\mathbf{a}_i \sim \pi_\theta^{old}$

    if $\mathbf{a}_i \notin \mathcal{B}_i$ **then**

      $w_i \leftarrow (1 - w_i^+)$

      $D \leftarrow D \cup (\mathbf{a}_i, R(\mathbf{a}_i), w_i)$

  Push D to training queue

until converge or early stop

repeat ▷ **for the learner**

  Get a batch D from training queue

  for $(\mathbf{a}_i, R(\mathbf{a}_i), w_i) \in D$ **do**

    $\mathrm{d}\theta \leftarrow \mathrm{d}\theta + w_i\, R(\mathbf{a}_i)\, \nabla \log \pi_\theta(\mathbf{a}_i)$

  update θ using $\mathrm{d}\theta$

  $\pi_\theta^{old} \leftarrow \pi_\theta$ ▷ once every M batches

until converge or early stop

Output: final parameters θ

Results :

	WikiTable	WikiSQL
REINFORCE	< 10	< 10
MML (Soft EM)	39.7 ± 0.3	70.7 ± 0.1
Hard EM	39.3 ± 0.6	70.2 ± 0.3
IML	36.8 ± 0.5	70.1 ± 0.2
MAPO	$\mathbf{42.3} \pm 0.3$	$\mathbf{72.2} \pm 0.2$
MAPO w/o SE	< 10	< 10
MAPO w/o MWC	< 10	< 10

Performance Comparison with other models

WIKITABLES :

	E.S.	Dev.	Test
Pasupat & Liang (2015) [39]	-	37.0	37.1
Neelakantan *et al.* (2017) [34]	1	34.1	34.2
Neelakantan *et al.* (2017) [34]	15	37.5	37.7
Haug *et al.* (2017) [18]	1	-	34.8
Haug *et al.* (2017) [18]	15	-	38.7
Zhang *et al.* (2017) [67]	-	40.4	43.7
MAPO	1	**42.7**	43.8
MAPO (mean of 5 runs)	-	42.3	43.1
MAPO (std of 5 runs)	-	0.3	0.5
MAPO (ensembled)	10	-	**46.3**

Table 3: Results on WIKITABLEQUESTIONS. E.S. is the ensemble size, when applicable.

Performance Comparison with other models

WIKISQL :

Fully supervised	Dev.	Test
Zhong *et al.* (2017) [68]	60.8	59.4
Wang *et al.* (2017) [56]	67.1	66.8
Xu *et al.* (2017) [61]	69.8	68.0
Huang *et al.* (2018) [22]	68.3	68.0
Yu *et al.* (2018) [63]	74.5	73.5
Sun *et al.* (2018) [54]	75.1	74.6
Dong & Lapata (2018) [14]	**79.0**	**78.5**
Weakly supervised	Dev.	Test
MAPO	**72.4**	72.6
MAPO (mean of 5 runs)	72.2	72.1
MAPO (std of 5 runs)	0.2	0.3
MAPO (ensemble of 10)	-	**74.9**

Table 4: Results on WIKISQL. Unlike other methods, MAPO only uses weak supervision.

Possible Extension and future work :

- Incorporate strong supervision to improve cold start problem (RAJAS)
- Use predefined SQL templates and learn to fill them ie - do an easier task (ATISHYA)
- Use simpler language with less redundant functionality - (PRATYUSH)
- Use this general technique for various other tasks (SIDDHANT,PAWAN)
- Add a language model for SQL queries (SARANSH)

Thank You