Memory Augmented Policy Optimization (MAPO) for Program Synthesis and Semantic Parsing

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

NORTHWESTERN UNIVERSITY

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

12,467

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

(filter$_{in}$ rows ['saskatoon'] r.city)
(filter$_{in}$ rows ['los angeles'] r.city)
(diff v1 v0 r.passengers)

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

12,467

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

$(\texttt{filter}_{\texttt{in}}$ `rows ['saskatoon'] r.city)`
$(\texttt{filter}_{\texttt{in}}$ `rows ['los angeles'] r.city)`
`(diff v1 v0 r.passengers)`

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

12,467

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

```
(filter_in rows ['saskatoon'] r.city)
(filter_in rows ['los angeles'] r.city)
(diff v1 v0 r.passengers)
```

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

12,467

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

```
(filter_in rows ['saskatoon'] r.city)
(filter_in rows ['los angeles'] r.city)
(diff v1 v0 r.passengers)
```

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

12,467

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

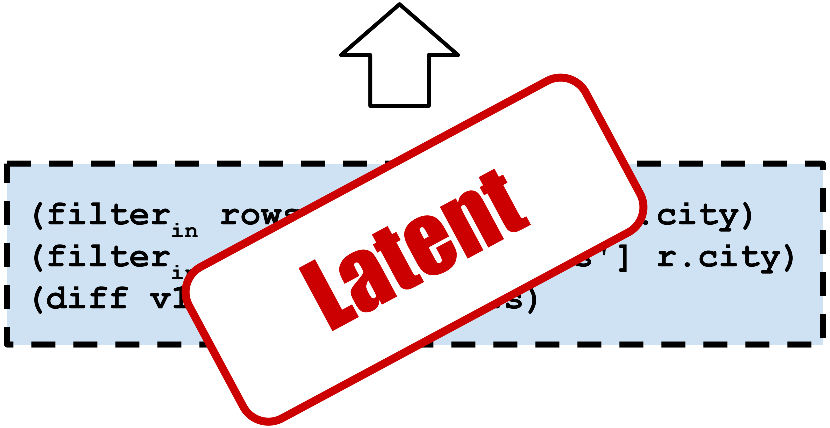
```
(filter_in rows ['saskatoon'] r.city)
(filter_in rows ['los angeles'] r.city)
(diff v1 v0 r.passengers)
```

Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

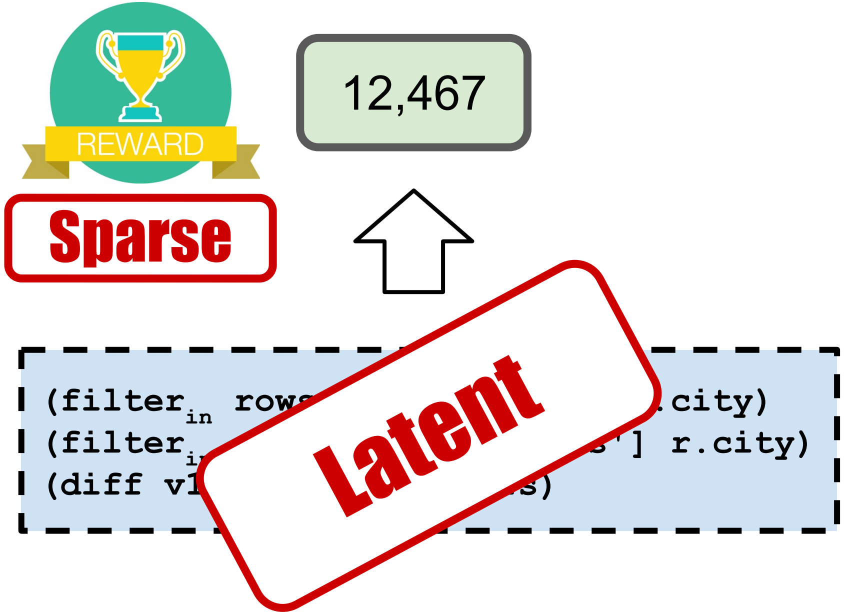
12,467

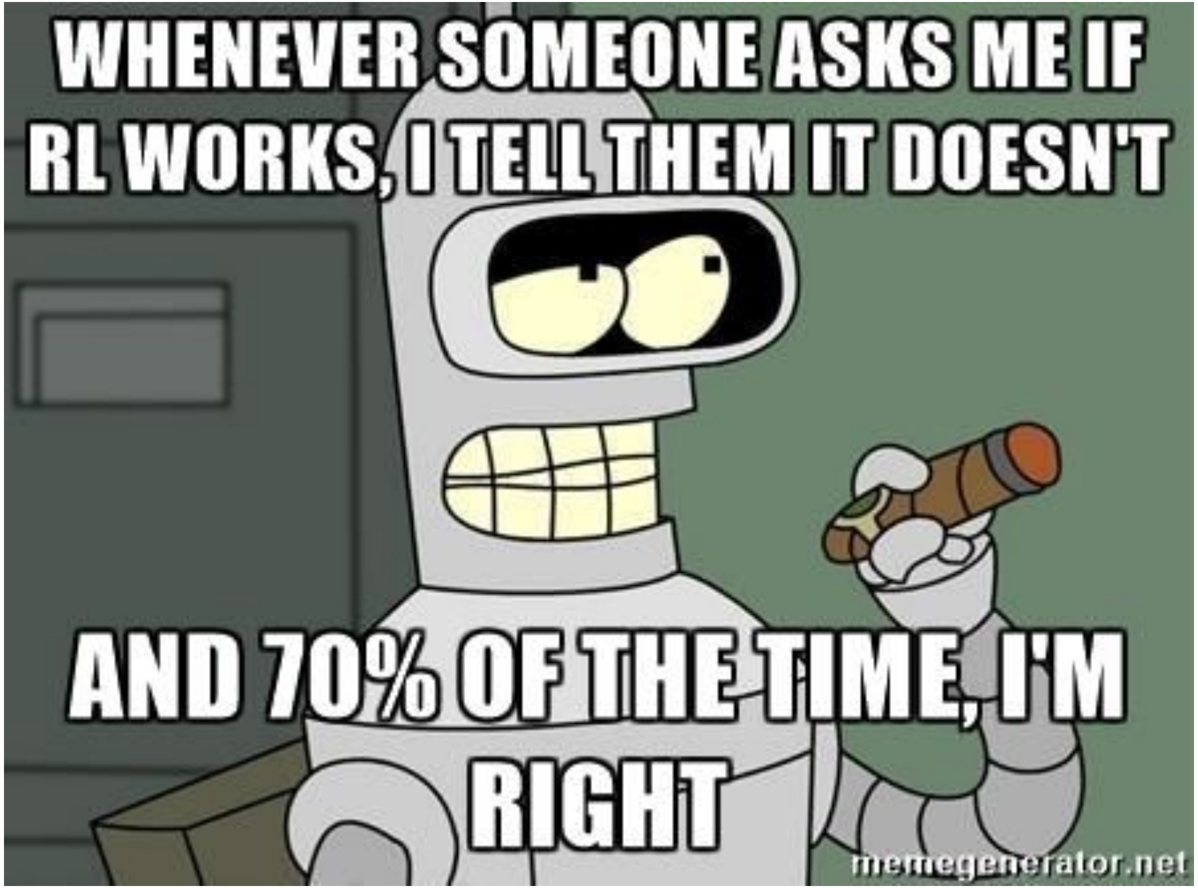


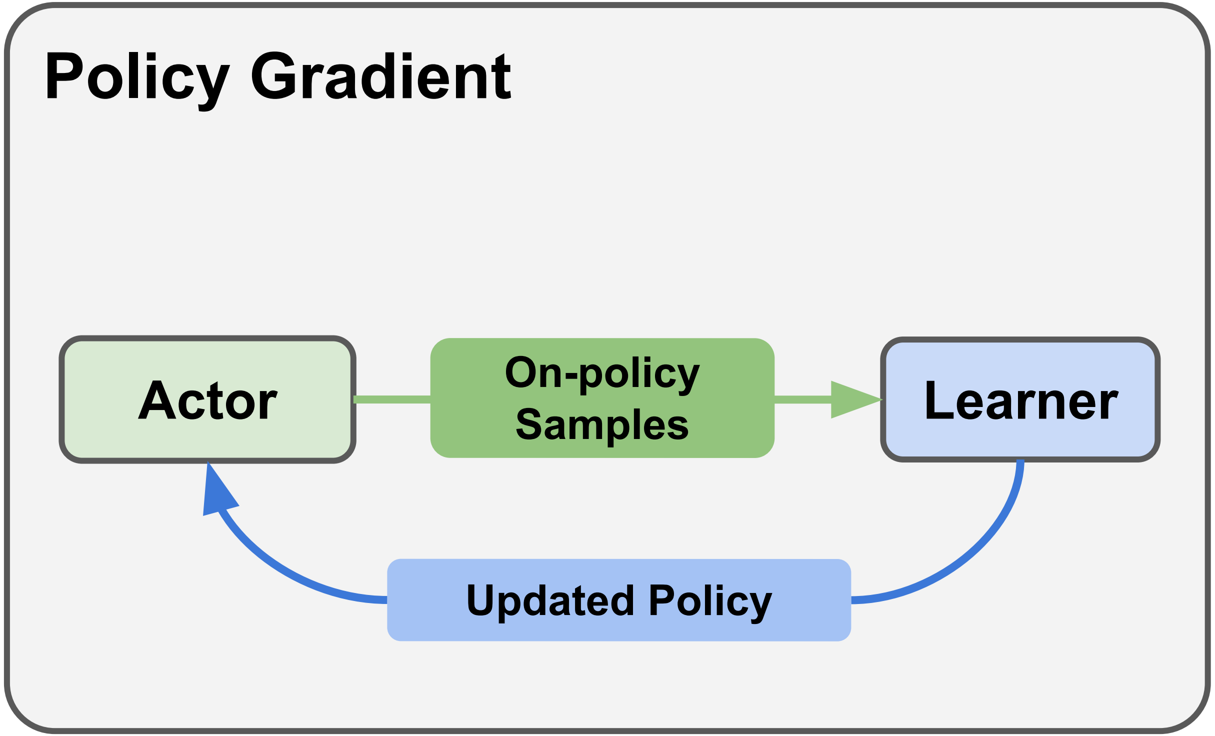
Program Synthesis / Semantic Parsing

how many more passengers flew to los angeles than to saskatoon?

Rank	City	Passengers	Ranking	Airline
1	United States, Los Angeles	14,749		Alaska Airlines
2	United States, Houston	5,465		United Express
3	Canada, Calgary	3,761		Air Transat, WestJet
4	Canada, Saskatoon	2,282	4	
5	Canada, Vancouver	2,103		Air Transat
6	United States, Phoenix	1,829	1	US Airways
7	Canada, Toronto	1,202	1	Air Transat, CanJet
8	Canada, Edmonton	110		
9	United States, Oakland	107		

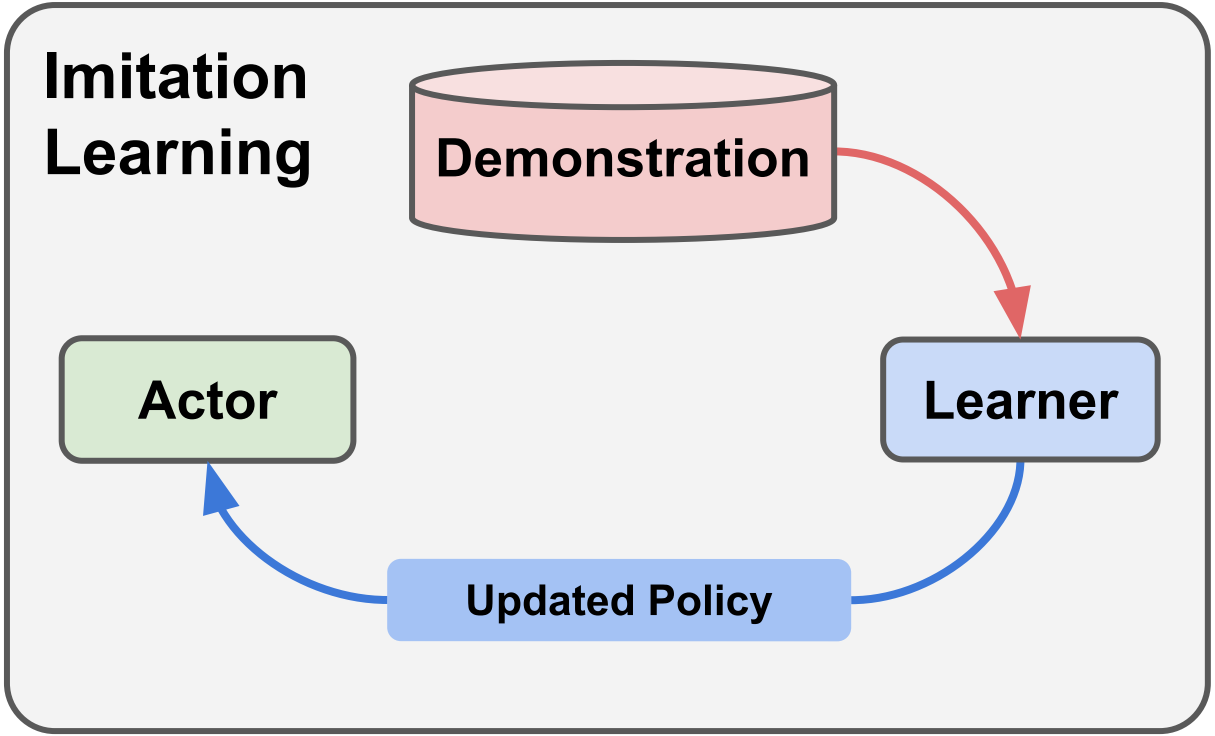



WHENEVER SOMEONE ASKS ME IF RL WORKS, I TELL THEM IT DOESN'T
AND 70% OF THE TIME, I'M RIGHT
memegenerator.net



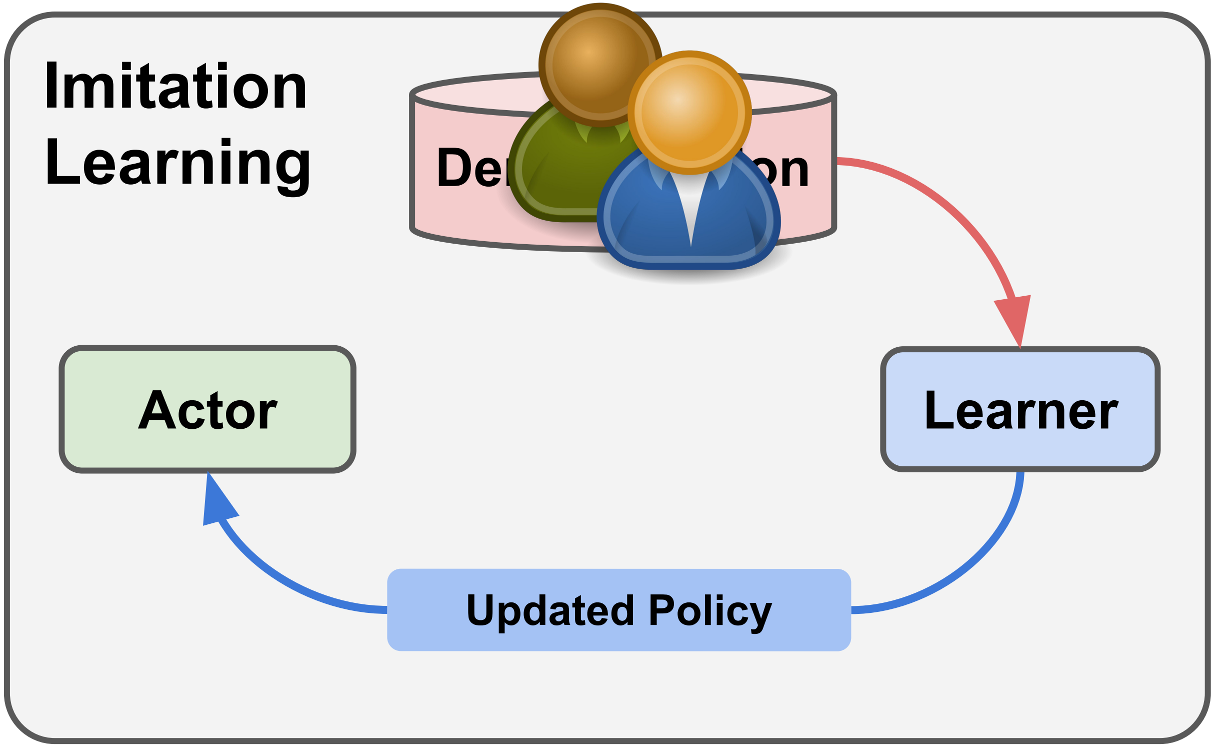
Unbiased => optimal solution

High variance => slow training

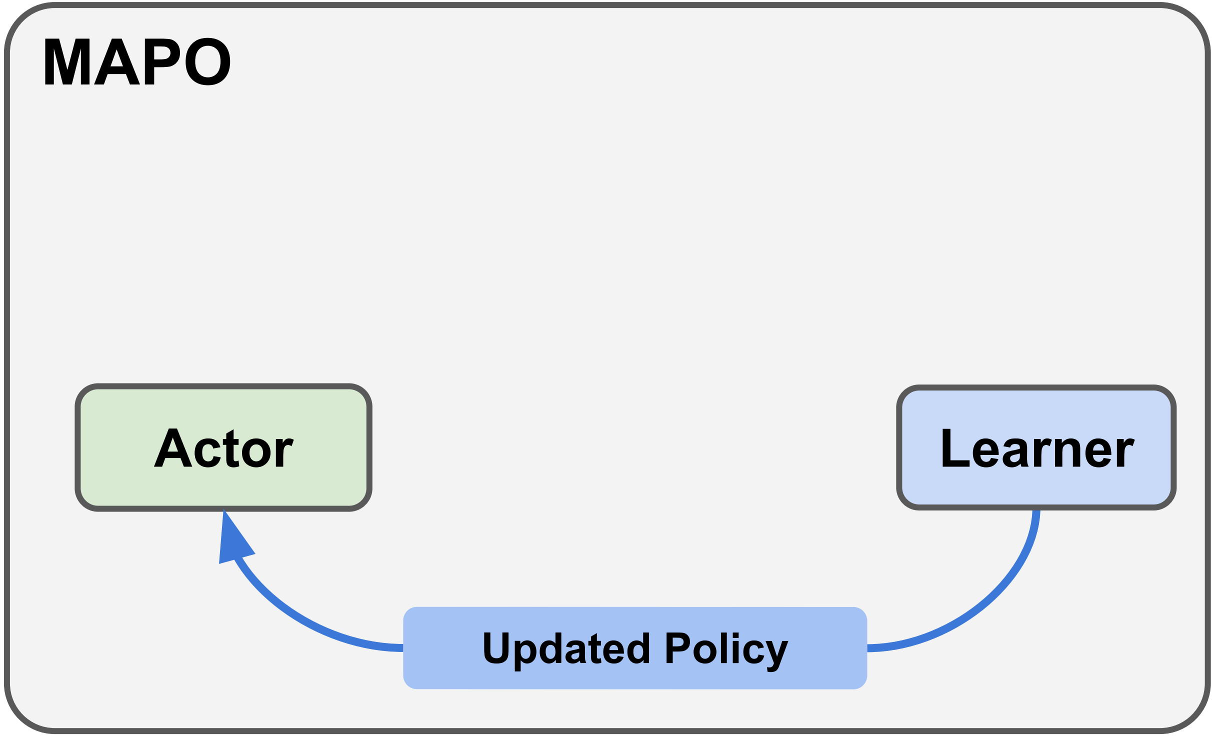


Biased => suboptimal solution

Low variance => fast training

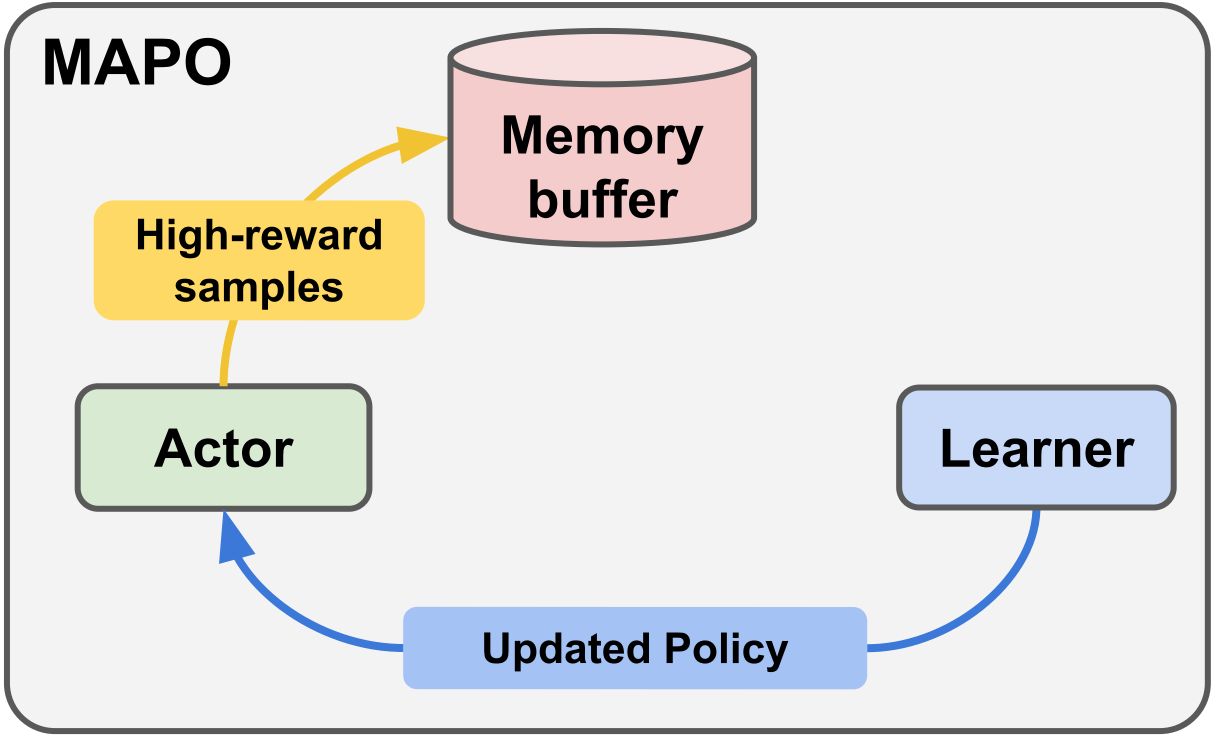


- ❌ Biased => suboptimal solution
- ✅ Low variance => fast training
- ❌ Requires human supervision

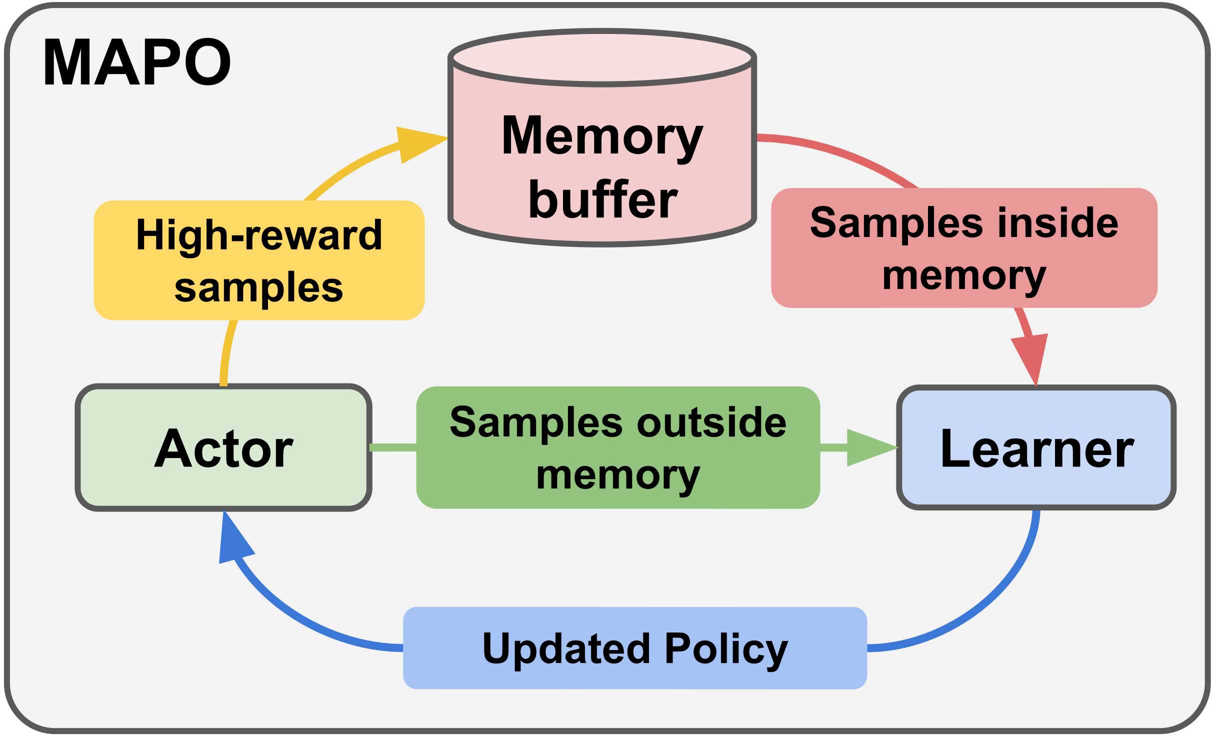


Unbiased => optimal solution

Low variance => fast training

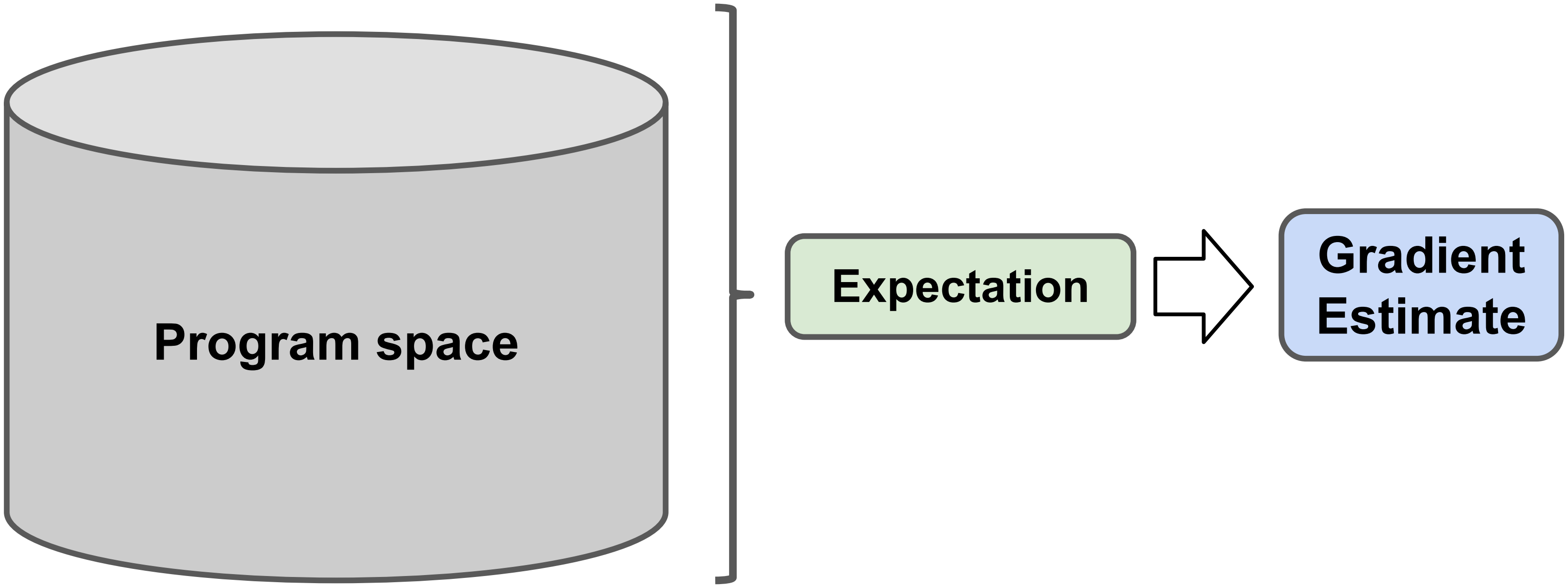


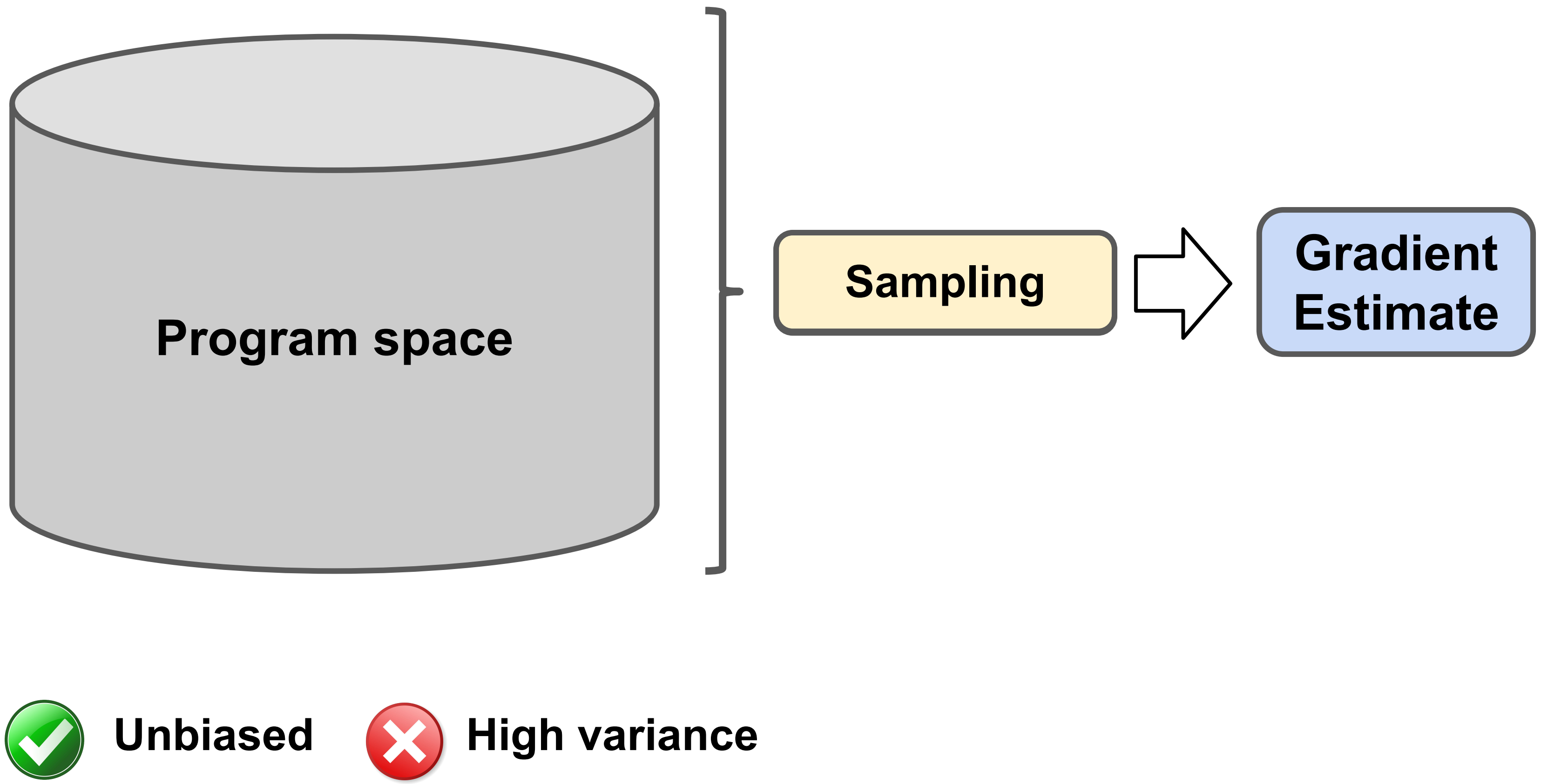
- Unbiased => optimal solution
- Low variance => fast training



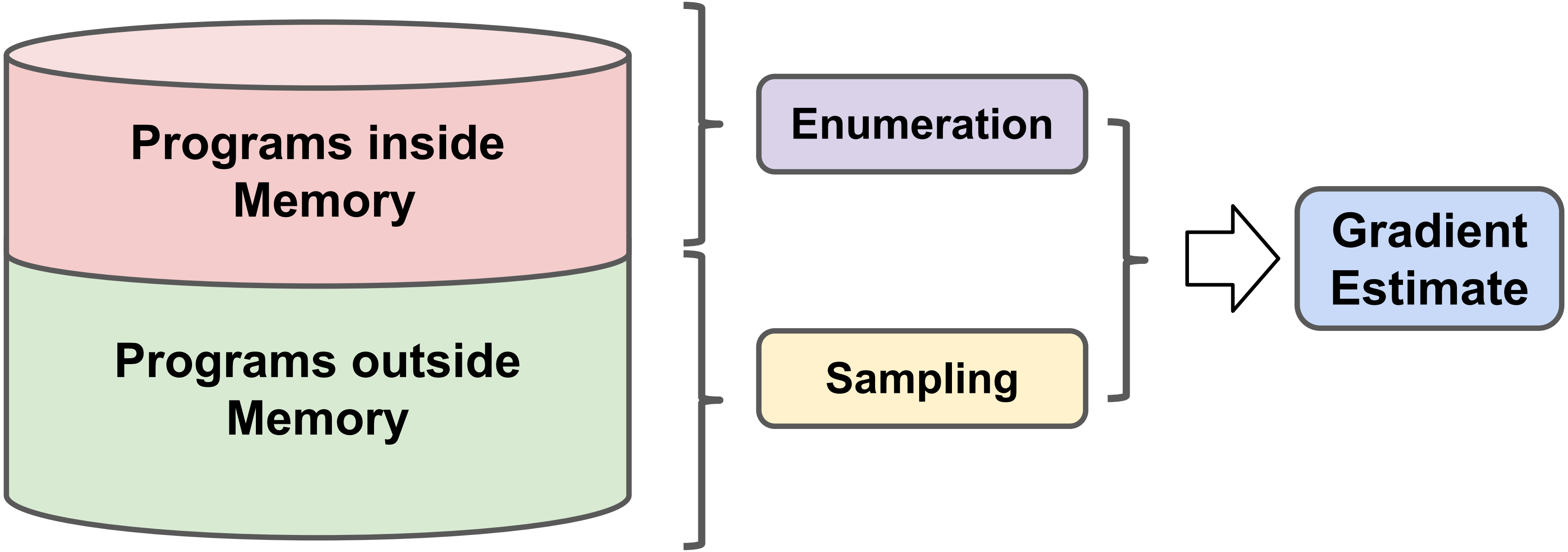
Unbiased => optimal solution

Low variance => fast training


Program space
Expectation
Gradient
Estimate

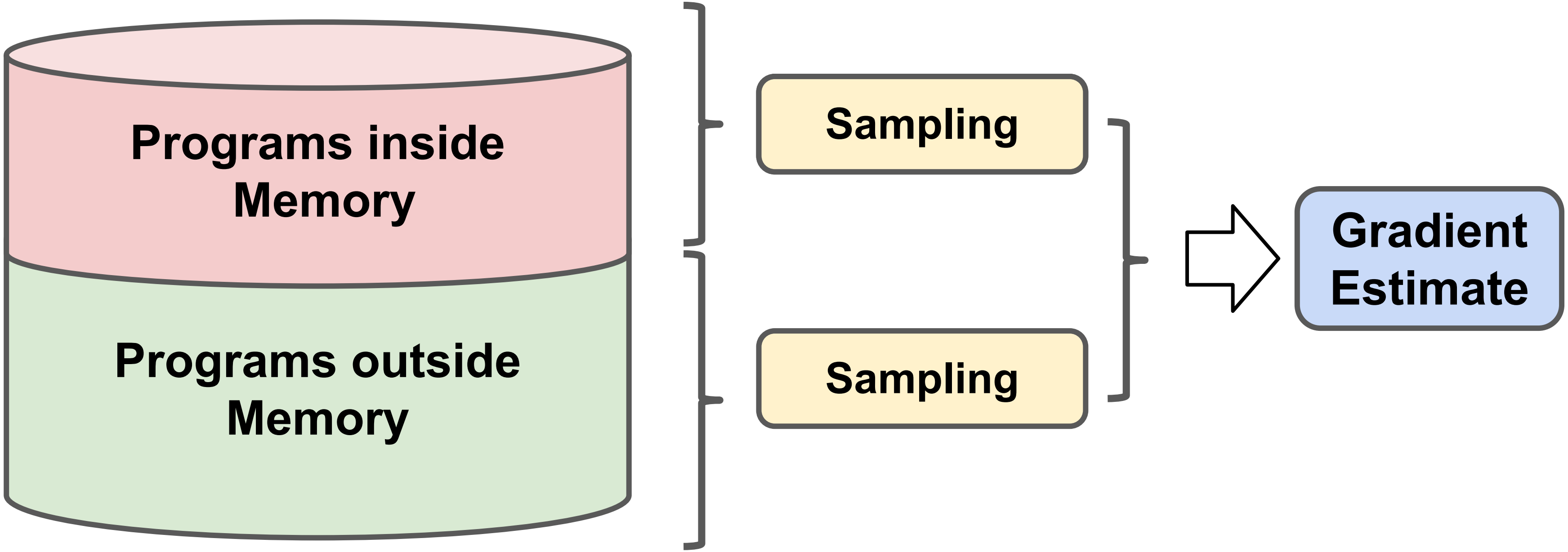

Program space
Sampling
Gradient Estimate
Unbiased
High variance

MAPO



Unbiased   Sampling from a smaller space => variance reduction

MAPO



✅ Unbiased ✅ Stratified sampling => variance reduction

MAPO

$$O(\pi) = \sum_{\vec{a} \in \mathcal{A}} \pi(\vec{a}) R(\vec{a})$$

($\vec{a}$ = a program)
($R(\vec{a})$ = correct or not)

MAPO

$$O(\pi) = \sum_{\vec{a} \in \mathcal{A}} \pi(\vec{a}) R(\vec{a})$$

($\vec{a}$ = a program)
($R(\vec{a})$ = correct or not)

$$O(\pi) = \underbrace{\sum_{\vec{a}_i \in \mathcal{B}} \pi(\vec{a}_i) r_i}_{\text{Expectation inside } \mathcal{B}} + \underbrace{\sum_{\vec{a} \notin \mathcal{B}} \pi(\vec{a}) R(\vec{a})}_{\text{Expectation outside } \mathcal{B}}$$

$$\mathcal{B} \equiv \{(\vec{a}_i, r_i)\}_{i=1}^{N}$$

WikiTableQuestions: first SOTA using RL

	E.S.	Dev.	Test
Pasupat & Liang (2015)	-	37.0	37.1
Neelakantan *et al.* (2017)	1	34.1	34.2
Neelakantan *et al.* (2017)	15	37.5	37.7
Haug *et al.* (2017)	1	-	34.8
Haug *et al.* (2017)	15	-	38.7
Zhang *et al.* (2017)	-	40.4	43.7

WikiTableQuestions: first SOTA using RL

	E.S.	Dev.	Test
Pasupat & Liang (2015)	-	37.0	37.1
Neelakantan *et al.* (2017)	1	34.1	34.2
Neelakantan *et al.* (2017)	15	37.5	37.7
Haug *et al.* (2017)	1	-	34.8
Haug *et al.* (2017)	15	-	38.7
Zhang *et al.* (2017)	-	40.4	43.7
MAPO	1	$\mathbf{42.4} \pm 0.5$	$\mathbf{43.2} \pm 0.5$
MAPO (ensembled)	10	-	**46.6**

WikiSQL: strong vs. weak supervision!

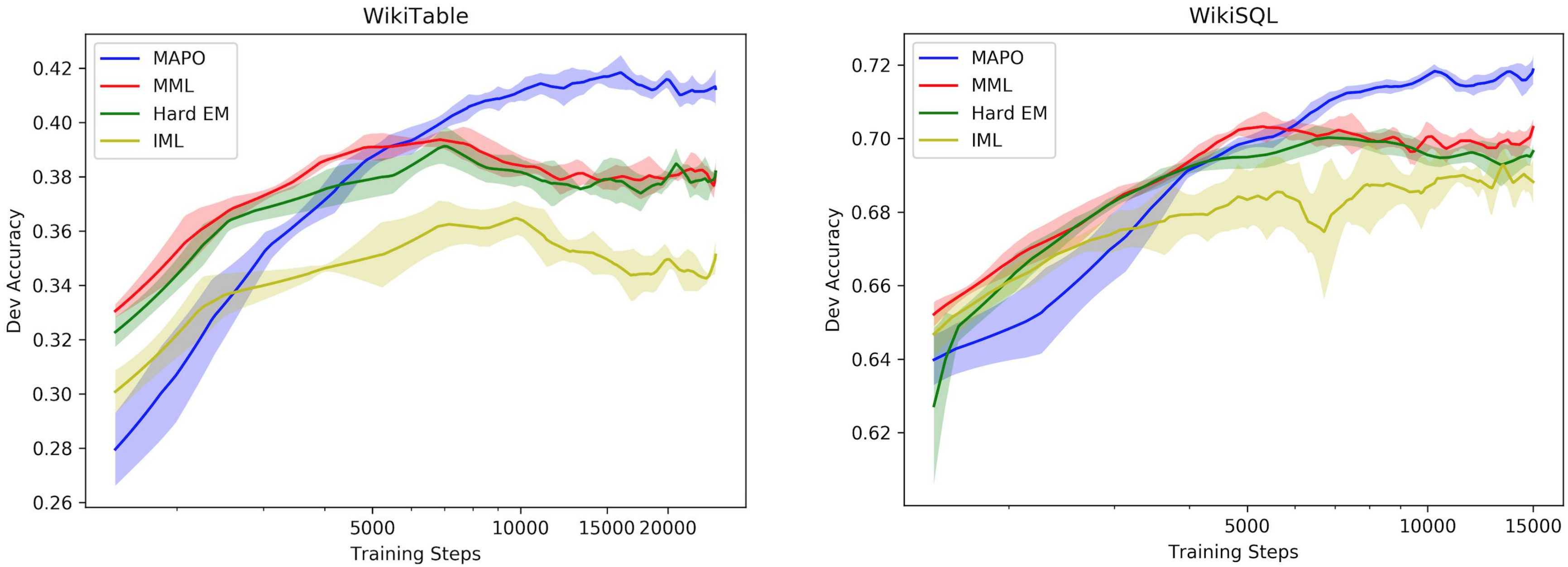
Fully supervised	Dev.	Test
Zhong *et al.* (2017)	60.8	59.4
Wang *et al.* (2017)	67.1	66.8
Xu *et al.* (2017)	69.8	68.0
Huang *et al.* (2018)	68.3	68.0
Yu *et al.* (2018)	74.5	73.5
Sun *et al.* (2018)	75.1	74.6
Dong & Lapata (2018)	**79.0**	**78.5**

Strong supervision

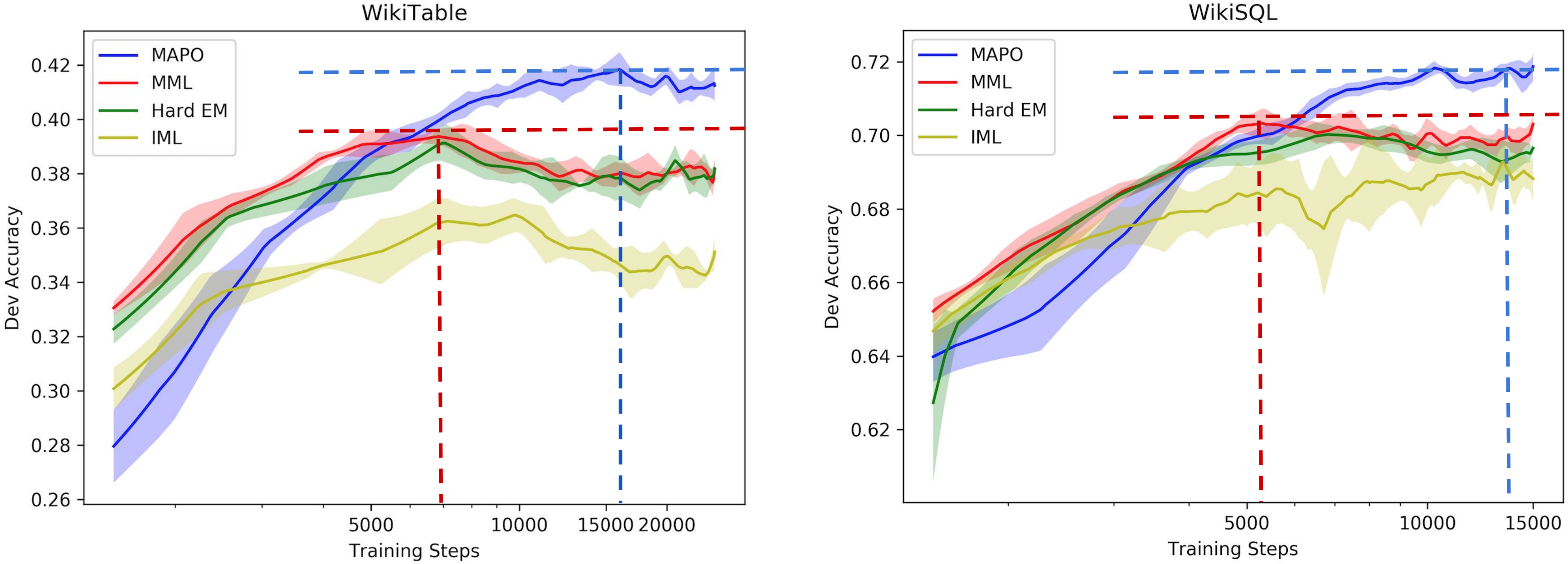
WikiSQL: strong vs. weak supervision!

Fully supervised		Dev.	Test
Zhong *et al.* (2017)	Strong supervision	60.8	59.4
Wang *et al.* (2017)		67.1	66.8
Xu *et al.* (2017)		69.8	68.0
Huang *et al.* (2018)		68.3	68.0
Yu *et al.* (2018)		74.5	73.5
Sun *et al.* (2018)		75.1	74.6
Dong & Lapata (2018)		**79.0**	**78.5**

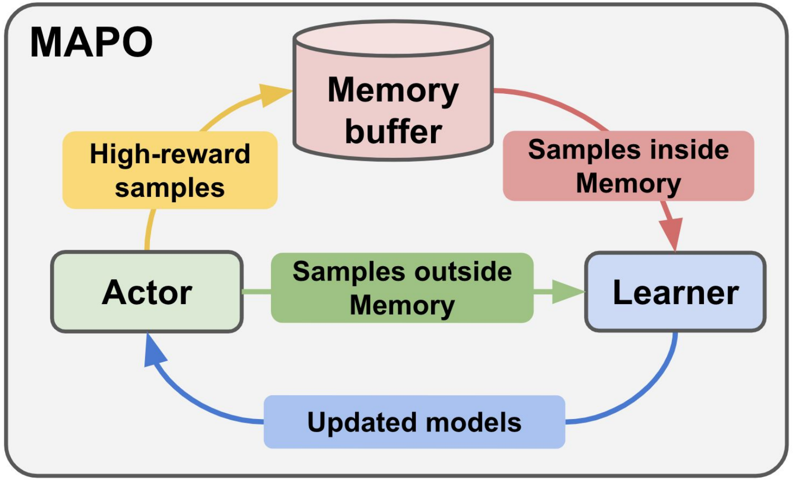
Weakly supervised	Dev.	Test
MAPO	$\mathbf{71.6} \pm 0.6$	$\mathbf{71.8} \pm 0.4$
MAPO (ensemble of 5)	-	**74.9**



- **MAPO converges slower than iterative maximum likelihood, but reaches a better solution.**
- **REINFORCE doesn’t make much progress (<10% accuracy).**



- **MAPO converges slower than maximum likelihood training, but reaches a better solution.**
- **REINFORCE doesn't make much progress (<10% accuracy).**



An efficient policy optimization method for learning to generate sequences from sparse rewards.

https://github.com/crazydonkey200/neural-symbolic-machines

https://arxiv.org/abs/1807.02322

http://crazydonkey200.github.io/

Poster: Room 517 AB #137