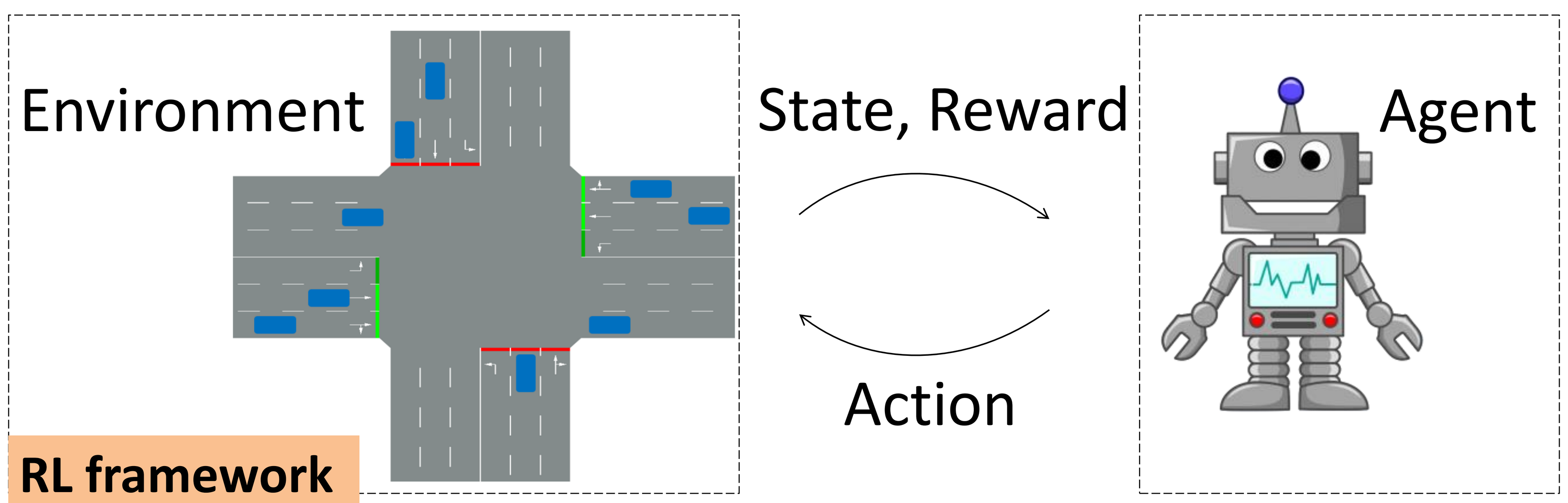
MetaLight: Value-based Meta-reinforcement Learning for Traffic Signal Control

Xinshi Zang, Huaxiu Yao, Guanjie Zheng, Nan Xu, Kai Xu, Zhenhui Li

Shanghai Jiao Tong University, Pennsylvania State University, Shanghai Tianrang Intelligent Technology Co., Ltd

Introduction

- Reinforcement Learning has made traffic signal control more intelligent.

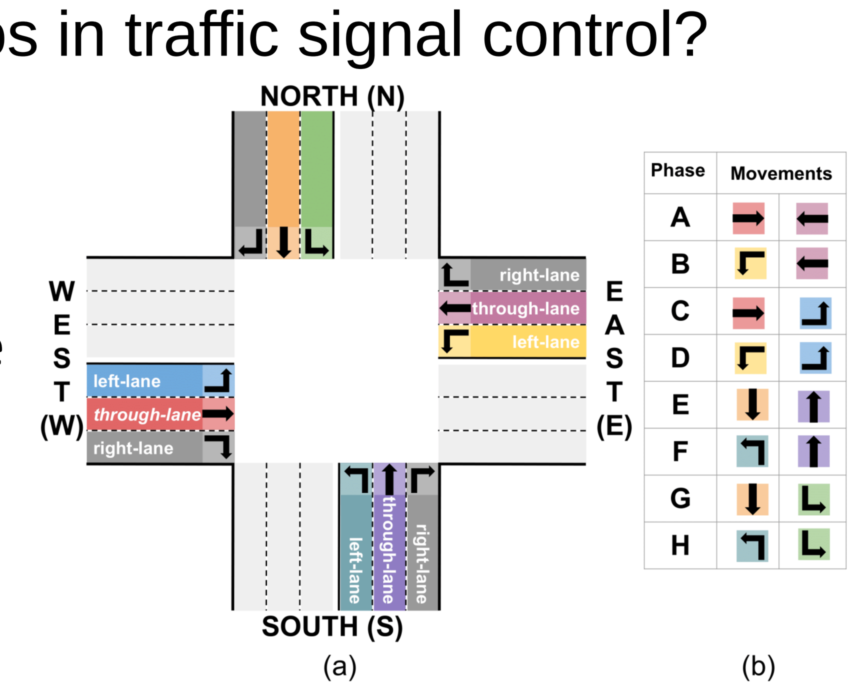


- However, the cost of computational resources and learning time is unacceptable in the complicated real world.

- Therefore, how about leveraging the common knowledge shared in all intersection scenarios (Meta-RL)?

Challenges

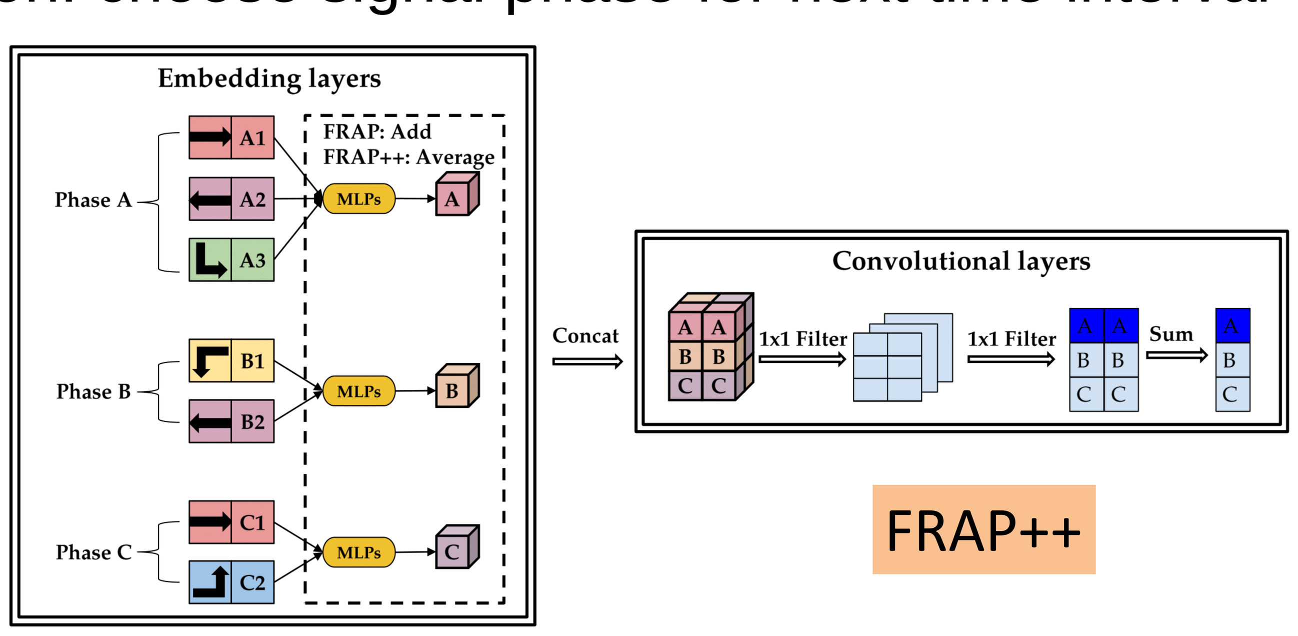
1. How to learn and adapt to the complicated and heterogeneous scenarios in traffic signal control?
   - Traffic flow
   - Entering approach/lane
   - Phase Setting



2. How apply meta-learning on value-based reinforcement learning?

Framework

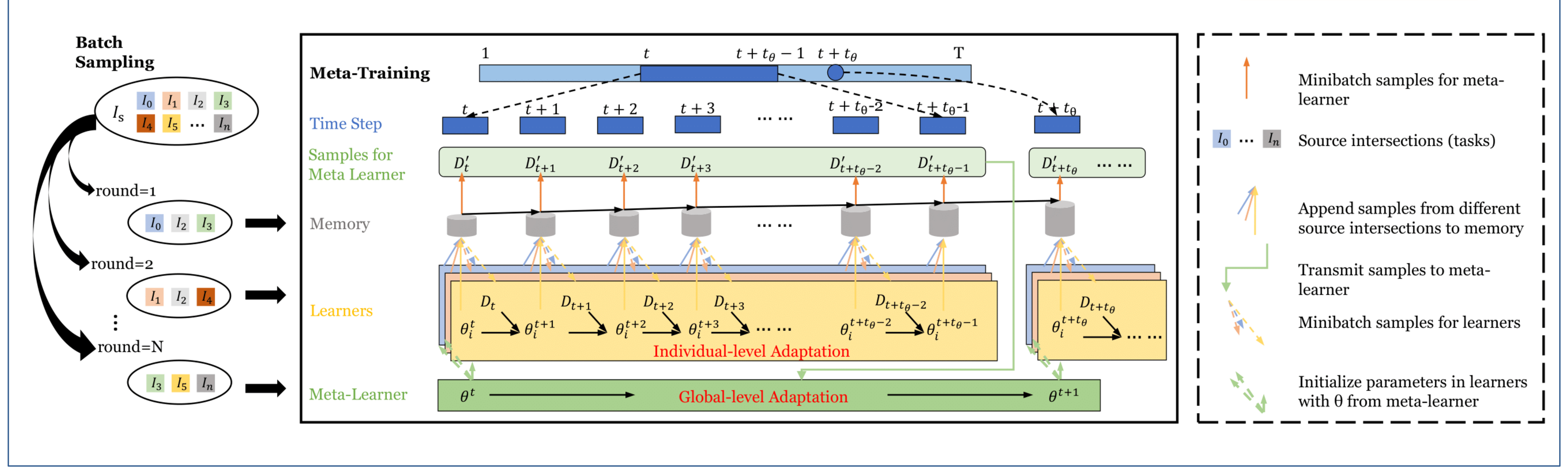
- State: number of vehicles on each lane
- Reward: average queue length
- Action: choose signal phase for next time interval



Framework

Value-based Meta-RL workflow based on the gradient-based meta-learning algorithm (MAML)

MetaLight



Experiments

Table: Performances on Task-1, 2, 3. Travel time is reported

Homogeneous

Phase Setting	4a	4b	6a	6c	6e	8
Random	102.71	292.51	90.41	461.78	105.49	73.62
Pretrained	82.87	**191.83**	85.47	200.06	111.94	67.88
MAML	82.95	**191.53**	161.41	404.04	132.26	77.07
MetaLight	**74.67**	199.55	**78.92**	**195.92**	**98.58**	**66.93**
Improvement	9.89%	\	7.66%	2.07%	6.56%	1.41%

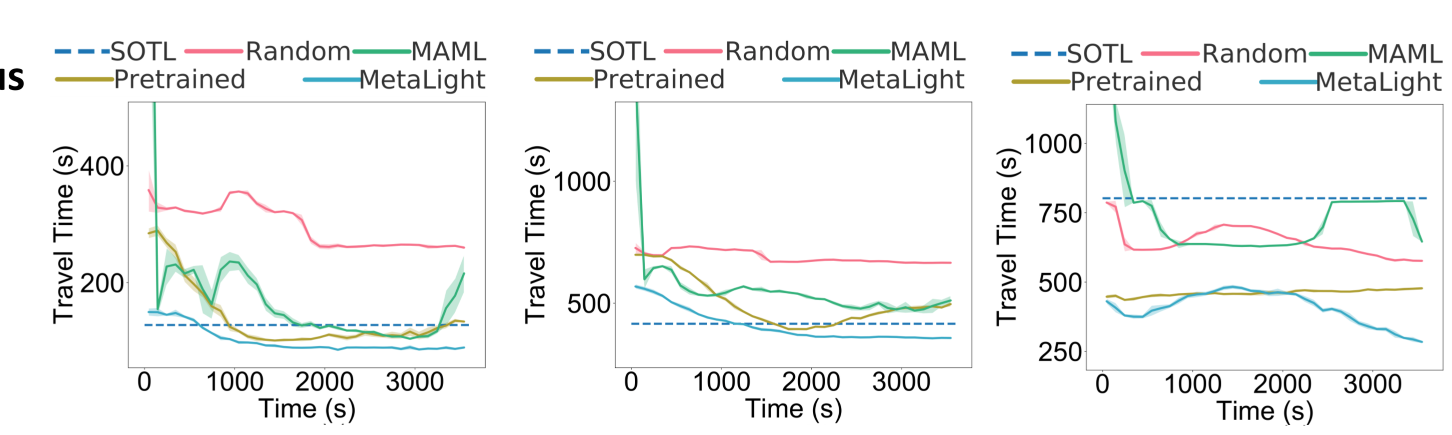
Heterogeneous

Phase Setting	4c	4d	6b	6d	6f
Random	254.70	662.85	298.53	570.55	474.20
Pretrained	95.94	385.74	233.64	430.74	307.98
MAML	101.41	440.65	369.82	614.09	345.46
MetaLight	**81.07**	**352.83**	**172.91**	**273.58**	**226.82**
Improvement	15.50%	8.53%	25.99%	36.49%	26.35%

Different Cities

City	Homogeneous			Heterogeneous		
	JN	AT	LA	JN	AT	LA
Random	451.88	379.16	262.23	363.59	602.60	684.15
Pretrained	128.20	186.86	104.59	156.04	351.39	331.75
MAML	173.13	301.29	135.11	335.81	618.84	393.58
MetaLight	**95.01**	**161.37**	**77.23**	**137.02**	**310.39**	**308.71**
Improvement	25.89%	13.64%	26.16%	10.17%	11.67%	6.94%

Heterogeneous



(a) Phase setting: 4c (b) Phase setting: 4d (c) Phase setting: 6d

Figure: Meta-testing curves for Task-2

Acknowledgements

The work was supported in part by NSF awards #1652525 and #1618448. The views and conclusions contained in this paper are those of the authors and should not be interpreted as representing any funding agencies

References

[KDD'18] Wei et al., IntelliLight: A Reinforcement Learning Approach for Intelligent Traffic Light Control
[CIKM'19] Zheng et al, Learning Traffic Signal Control from Demonstrations
[ICML'17] Finn et al, Model-agnostic meta-learning for fast adaptation of deep networks
[arXiv:1904.08117] Wei et al., A Survey on Traffic Signal Control Methods
[arXiv:1905.04716] Zheng et al., Diagnosing Reinforcement Learning for Traffic Signal Control

Try to find datasets, code, demo and more related researches?
Just scan QR code on the right or visit
https://traffic-signal-control.github.io