
Reinforcement Learning Benchmarks for Traffic Signal Control

James Ault

Texas A&M University
College Station, TX
jault@tamu.edu

Guni Sharon

Texas A&M University
College Station, TX
guni@tamu.edu

Abstract

We propose a toolkit for developing and comparing reinforcement learning (RL)-based traffic signal controllers. The toolkit includes implementation of state-of-the-art deep-RL algorithms for signal control along with benchmark control problems that are based on realistic traffic scenarios. Importantly, the toolkit allows a first-of-its-kind comparison between state-of-the-art RL-based signal controllers while providing benchmarks for future comparisons. Consequently, we compare and report the relative performance of current RL algorithms. The experimental results suggest that previous algorithms are not robust to varying sensing assumptions and non-stylized intersection layouts. When more realistic signal layouts and advanced sensing capabilities are considered, a distributed deep Q-learning approach is shown to outperform previously reported state-of-the-art algorithms in many cases.

A Appendix

A.1 License

RESCO is open source and free to use/modify under the GNU General Public License 3. The code is available on Github at [<https://github.com/Pi-Star-Lab/RESCO>]. Included benchmark scenarios are distributed with their licenses. Cologne based scenarios are under creative commons BY-NC-SA and Ingolstadt based, with the GNU General Public License 3. All experiments can be reproduced from the source code, which includes all hyper-parameters and configuration.

A.2 Supplementary results

Figure 1 displays the full learning curves of the IPPO and FMA2C algorithms.

Table 1 shows the final, post-training performance for each RL algorithm on the benchmark control tasks. These results show the learning divergence for MPLight in the Cologne Corridor, Cologne Region, Ingolstadt Corridor, and Ingolstadt Region scenarios as compared to the best performance.

Table 2 gives the episode number in which the best performance from Table ?? were reached. It supports the conclusion that MPLight, followed by IDQN, have the best sample efficiency of the compared algorithms.

To produce Table 3, the standard deviation of each episode over the 5 random seeds is calculated for each metric. The average over all episodes is then displayed in Table 3 per algorithm, allowing for a notion of stability. FMA2C displays better stability than any of IDQN, IPPO, and MPLight.

Table 4 extends Table ?? by including standard deviations. Table ?? reports the best performing episode of each algorithm averaged over five random seeds.

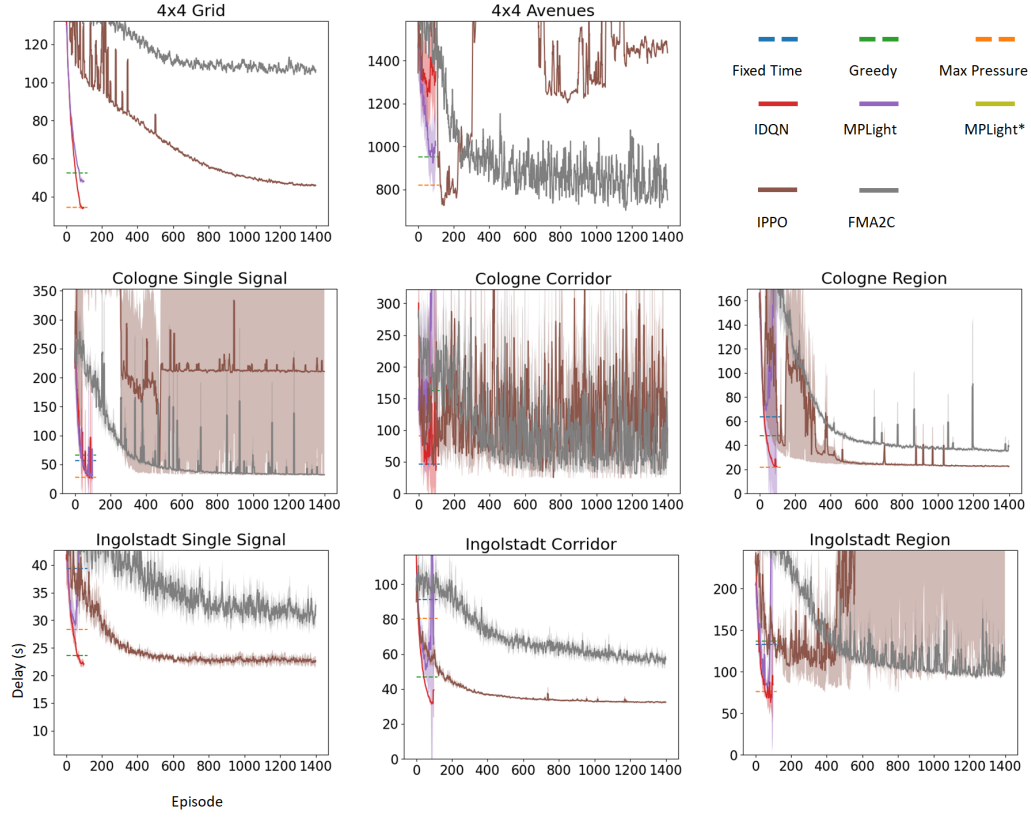


Figure 1: Full Learning curves over 5 random seeds

Table 5 provides the parameters of each algorithm, note that IDQN and MPLight share the same parameters.

Table 1: Final Performance on benchmark scenarios

<i>IDQN</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	21.48	31.19	61.11	26.05	24.55	22.06
Avg. Trip Time	35.29	68.69	198.72	43.59	59.0	86.02
Avg. Wait	3.93	8.71	21.36	7.98	8.5	5.46
Avg. Queue	0.43	0.67	0.82	2.09	0.87	0.38
<i>IPPO</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	21.85	32.0	369.34	209.49	58.98	21.62
Avg. Trip Time	35.68	69.38	509.28	224.56	93.7	85.71
Avg. Wait	4.2	8.86	315.31	193.66	43.5	5.19
Avg. Queue	0.49	0.74	5.45	21.86	5.73	0.36
<i>MPLight</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	*29.93	50.81	275.21	29.0	579.69	164.91
Avg. Trip Time	*42.82	79.95	406.27	46.17	595.36	225.14
Avg. Wait	*9.78	16.0	230.88	8.61	548.42	142.35
Avg. Queue	*0.65	1.35	4.69	2.49	23.77	5.31
<i>FMA2C</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	27.74	51.78	96.67	30.82	32.58	34.44
Avg. Trip Time	41.4	87.75	233.04	48.07	66.83	98.68
Avg. Wait	8.74	23.32	48.88	11.87	15.84	14.7
Avg. Queue	1.07	1.86	1.84	3.2	2.01	1.03

Table 2: Episodes before best performance

<i>Algorithm</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
IDQN	94	93	85	94	82	98
IPPO	1052	1373	386	475	744	1394
MPLight	*76	59	71	83	34	35
FMA2C	1361	1339	1372	1230	1375	1380

Table 3: Average standard deviation over all trials on benchmark scenarios

<i>IDQN</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	1.56	3.27	17.82	37.03	77.93	6.14
Avg. Trip Time	1.22	2.59	18.4	31.34	75.68	5.64
Avg. Wait	0.91	2.08	15.83	30.28	77.35	4.11
Avg. Queue	0.11	0.1	0.36	3.25	2.71	0.26
<i>IPPO</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	1.14	0.88	196.31	280.04	72.07	13.27
Avg. Trip Time	0.9	0.74	191.44	277.68	71.8	13.95
Avg. Wait	0.68	0.56	190.07	282.48	73.2	12.41
Avg. Queue	0.06	0.03	1.84	16.46	4.2	0.75
<i>MPLight</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	*3.5	32.03	70.11	25.54	189.9	153.38
Avg. Trip Time	*2.64	22.43	68.51	20.78	183.75	146.09
Avg. Wait	*2.13	21.61	67.21	19.97	185.82	141.03
Avg. Queue	*0.13	0.65	1.05	1.86	6.5	4.23
<i>FMA2C</i>	Ing. Single	Ing. Corr.	Ing. Reg.	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	3.27	4.34	12.49	11.67	45.28	5.35
Avg. Trip Time	2.4	3.54	12.84	9.37	43.48	4.94
Avg. Wait	2.09	2.68	11.01	8.82	44.43	4.02
Avg. Queue	0.19	0.1	0.23	1.06	1.67	0.23

Table 4: Performance on benchmark scenarios with standard deviations

<i>IDQN</i>	Ing. Single	Ing. Corr.	Ing. Reg.
Avg. Delay	21.48 \pm 0.56	31.19 \pm 0.97	59.64 \pm 2.13
Avg. Trip Time	35.29 \pm 0.48	68.69 \pm 0.72	197.23 \pm 2.18
Avg. Wait	3.93 \pm 0.25	8.71 \pm 0.56	20.19 \pm 1.48
Avg. Queue	0.43 \pm 0.01	0.67 \pm 0.03	0.8 \pm 0.05
<i>IPPO</i>	Ing. Single	Ing. Corr.	Ing. Reg.
Avg. Delay	19.85 \pm 0.21	30.7 \pm 0.98	67.65 \pm 19.49
Avg. Trip Time	34.19 \pm 0.26	68.34 \pm 0.77	205.44 \pm 8.4
Avg. Wait	3.21 \pm 0.28	8.2 \pm 0.52	26.45 \pm 15.12
Avg. Queue	0.39 \pm 0.02	0.71 \pm 0.0	1.15 \pm 0.56
<i>MPLight</i>	Ing. Single	Ing. Corr.	Ing. Reg.
Avg. Delay	28.31 \pm 0.8	48.21 \pm 4.24	78.16 \pm 2.06
Avg. Trip Time	41.07 \pm 1.01	76.58 \pm 1.42	215.72 \pm 2.21
Avg. Wait	8.27 \pm 0.97	15.05 \pm 1.73	34.57 \pm 1.78
Avg. Queue	0.61 \pm 0.03	1.34 \pm 0.06	1.48 \pm 0.04
<i>FMA2C</i>	Ing. Single	Ing. Corr.	Ing. Reg.
Avg. Delay	25.36 \pm 1.5	48.99 \pm 0.54	90.42 \pm 1.69
Avg. Trip Time	39.4 \pm 1.11	85.03 \pm 1.63	226.5 \pm 1.47
Avg. Wait	7.27 \pm 0.64	21.9 \pm 0.15	44.16 \pm 2.0
Avg. Queue	0.94 \pm 0.0	1.79 \pm 0.02	1.74 \pm 0.04
<i>IDQN</i>	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	26.05 \pm 0.57	23.99 \pm 1.11	22.06 \pm 0.36
Avg. Trip Time	43.59 \pm 0.52	59.0 \pm 0.87	86.02 \pm 0.39
Avg. Wait	7.98 \pm 0.35	8.5 \pm 0.59	5.46 \pm 0.2
Avg. Queue	2.09 \pm 0.1	0.87 \pm 0.02	0.38 \pm 0.02
<i>IPPO</i>	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	55.07 \pm 11.83	22.13 \pm 0.41	21.49 \pm 0.13
Avg. Trip Time	67.7 \pm 9.27	57.45 \pm 0.26	85.54 \pm 0.17
Avg. Wait	26.15 \pm 6.62	7.37 \pm 0.36	5.01 \pm 0.17
Avg. Queue	8.88 \pm 2.74	0.76 \pm 0.1	0.35 \pm 0.0
<i>MPLight</i>	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	28.74 \pm 1.67	83.65 \pm 27.96	60.42 \pm 20.17
Avg. Trip Time	45.85 \pm 1.14	102.3 \pm 21.53	123.93 \pm 20.43
Avg. Wait	8.61 \pm 0.65	46.25 \pm 19.27	30.34 \pm 15.48
Avg. Queue	2.45 \pm 0.24	5.4 \pm 1.94	2.33 \pm 0.97
<i>FMA2C</i>	Col. Single	Col. Corr.	Col. Reg.
Avg. Delay	30.12 \pm 0.3	25.37 \pm 0.17	33.28 \pm 0.49
Avg. Trip Time	47.31 \pm 0.17	61.68 \pm 0.21	97.53 \pm 0.43
Avg. Wait	11.23 \pm 0.01	11.3 \pm 0.03	14.19 \pm 0.88
Avg. Queue	3.11 \pm 0.05	1.68 \pm 0.04	0.98 \pm 0.06

Table 5: Algorithm Parameters

IDQN			
Optimizer: Adam Step Size: $1e-3$ Target Update Interval: 500 Steps	Adam β_1 : 0.9 Replay Size: 10,000	Adam β_2 : 0.999 Minibatch Size: 32	Adam ϵ : $1e-8$ γ : 0.99
MPLight			
Optimizer: Adam Step Size: $1e-3$ Target Update Interval: 500 Steps	Adam β_1 : 0.9 Replay Size: 10,000	Adam β_2 : 0.999 Minibatch Size: 32	Adam ϵ : $1e-8$ γ : 0.99
FMA2C			
Optimizer: RMSProp Entropy Coef: $1e-3$ Batch Size: 120	RMSProp α : 0.99 Entropy Ratio: 0.5 Reward Norm 2,000	RMSProp ϵ : $1e-5$ γ : 0.96 Reward Clip: 2	Step Size: $2.5e-4$ Value Coef: 0.5
IPPO			
Optimizer: Adam Step Size: $2.5e-4$ Update Interval: 1024 ϵ Clip: 0.1	Adam β_1 : 0.9 γ : 0.99 Minibatch Size: 256	Adam β_2 : 0.999 Entropy Coef: $1e-3$ Epochs 4	Adam ϵ : $1e-5$ Max L2 Norm: 0.5 λ : 0.95