



강화학습 기반 신호 최적화를 위한 State/Reward/Action 후보

- 목표
 - : 신호 최적화를 위한 시뮬레이터 요구사항 도출
 - : SALT 지원 계획 여부 파악

2020.07.23

내용

- 제공해야 하는 정보/기능 후보와 현재 상태
(강화학습기반 신호 최적화 선행 연구들에서 활용한 State/Reward/Action에 기반하여 정리함)
 - Get
 - State 후보와 시뮬레이터에서 정보 제공 여부
 - Reward 후보와 시뮬레이터에서 정보 제공 여부
 - Set
 - Action 후보와 시뮬레이터에서 기능 제공 여부
- 신호 최적화를 위한 Simulator 제공 기능 후보

GET : State 후보와 시뮬레이터에서 정보 제공 여부

0 : 현재, 1 : 1차년도, 2 : 2차년도, ...

State	Description	SUMO	SALT
<u>Queue Length</u> (of a lane)	차선에서 "Waiting"하고 있는 총 차량 수 (waiting 의 2 가지 정의 : speed < 0.1m/s or speed = 0)	O	0
<u>Waiting time</u> (of a vehicle)	차량이 "Waiting" 상태에 있는 시간 (2가지 정의 : 차량이 움직인 마지막 시간부터, or 차량이 도로에 들어온 시간부터) 차선/셀별 존재	O	0
Volume (of a lane)	차선의 차량 수 (= 기다리고 있는 차량 수 + 움직이고 있는 차량)	O	0
<u>Delay</u> (of a vehicle)	차량이 여행한 시간에서 예상했던 여정 시간을 뺀 값 (예상 여정 시간 = 거리 / 속도 제한)		
<u>speed</u> (of a vehicle)	차량이 얼마나 빠르게 여행하는지 (speed score = 차량의 실제 속도 / 속도 제한) SALT는 차선/셀별 존재 (일종의 space-mean speed?)	O	0
Phase duration (of the current phase)	현재 phase 가 얼마나 지속되었는지		
*Congestion (of a lane) (차량 길이 합 / 도로 길이)	혼잡 여부(0 or 1) 혹은 혼잡 수준 (= 차선의 차량 수 / 차선의 최대 허용 차량 수)		1
positions of vehicles	a matrix of vehicle positions, with '1' indicates the presence of a vehicle on a location, and '0' the absence of a vehicle on that location	O	
Phase	미리 정의된 Phase group에서 현재 phase 에 대한 인덱스	O	0
Elapsed time	차량이 도로에 들어와서 경과 시간	O	
Image	시뮬레이션 중 이미지		

* 다른 정보로부터 가공

GET : Reward 후보와 시뮬레이터에서 정보 제공 여부

0 : 현재, 1 : 1차년도, 2 : 2차년도, ...

Reward	Description	SUMO	SALT
<u>Queue Length</u>	모든 approaching lane 들의 대기 차량 수의 합 (queue length를 줄이는 것은 total travel time을 줄이는 것과 같다)	○	0
<u>Waiting time</u> (of a vehicle)	차량이 "waiting" 상태에 있어온 시간	○	0
<u>Change of delay</u>	총 누적 지연 시간의 차이 (two successive decision points) (현재 시스템에 있는 모든 차량의 누적 지연 시간의 합)	○	
<u>Speed</u>	도로의 모든 차량의 평균 속도	○	0
# of stops	도로의 모든 차량의 평균 멈춤 수	○	
Throughput	마지막 action 후 일정 시간이 지나는 동안 교차로를 통과하거나 도로 네트워크를 떠난 차량의 총 수	○	1
Frequency of signal change	특정 시간동안 신호가 변환 회수 (신호 변화가 많으면 차량이 교차로를 통과할 수 있는 "effective green time"이 작아진다)	○	
*Pressure of the intersection	모든 traffic movement에 대한 pressure의 합 (pressure 가 높다는 것은 진입/진출 차량의 불균형이 높다는 것을 의미)		
Space mean speed	어떤 공간에 존재하는 자동차 속도의 평균		0

* 다른 정보로부터 가공

SET : Action 후보와 시뮬레이터에서 제공 여부

0 : 현재, 1 : 1차년도, 2 : 2차년도, ...

Action	Description	SUMO	SALT
Set current phase duration	Agent가 사전 정의된 후보 시간 중 선택하여 현재 phase의 지속시간을 설정하는 방법을 학습한다.		1
Set cycle-based phase ratio	다음 Cycle에서 신호의 Phase Split Ratio the phase split ratio that the signal will set for the next cycle (Usually, the total cycle length is given, and the candidate phase ratio is pre-defined)		1
Keep or change the current phase (in a cycle-based phase sequence)	Agent 가 현재 phase를 유지할지, 다음 phase로 바꿀지를 지정한다.	O	1
Choose the next phase	변경할 phase를 결정 (다음에 취할 phase에 대한 인덱스) Agent는 신호가 주기적으로 변경될 것이라는 가정없이 변경될 phase를 선택하는 것을 학습한다.	O	1

신호 최적화를 위한 Simulator 제공 기능 후보

- 확정하기 어려움 & 다양한 정보를 줄 수 있으면 좋겠음

- GET

현재 1차년도 계산 가능 미지원

- 도로망/교차로별
 - 처리량(throughput, 교차로/네트워크 통과 차량)
 - 존재하는 차량 수
 - Pressure (혹은 각 approach에 대한 진입/진출 차량 수)
- 각 차선별
 - 대기열 길이, 차량 수(=대기 + 이동)
 - 차량들의 대기 시간(차량별이 아닌 차선/셀별 평균 대기 시간)
 - 차량들의 속도(차량별이 아닌 차선/셀별 평균 속도)
 - 혼잡 수준 (= [차량 길이 합 / 도로 길이] 혹은 [차선의 차량수/최대 허용 차량수])
- 각 차량별
 - 지연시간, 대기 시간, 속도, 멈춤 회수
- Phase
 - 현재 Phase, 신호 변경 회수

- SET

- 현재 phase의 지속 시간 설정/변경(Set current phase duration)
- 다음 번 Cycle의 Phase ratio 설정 (Set cycle-based phase ratio)
- Phase 변경
 - Cycle의 다음 Phase로 바꿈(Keep or change the current phase)
 - 인자로 받은 신호 Phase로 변경(Choose the next phase)