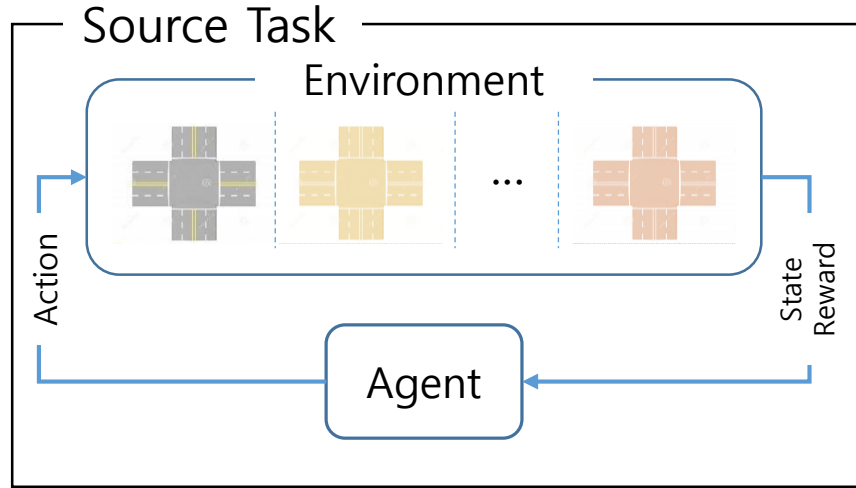

Multi-Input Multi-Output Model

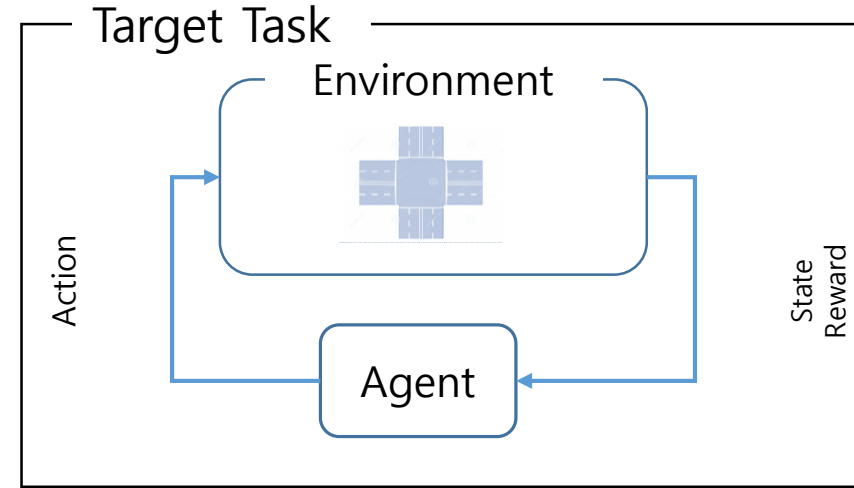
2023. 5. 3.
이용진

Training



N 개의 독립된 단일 교차로

Test



N+1번째 유사 교차로

Augmented State

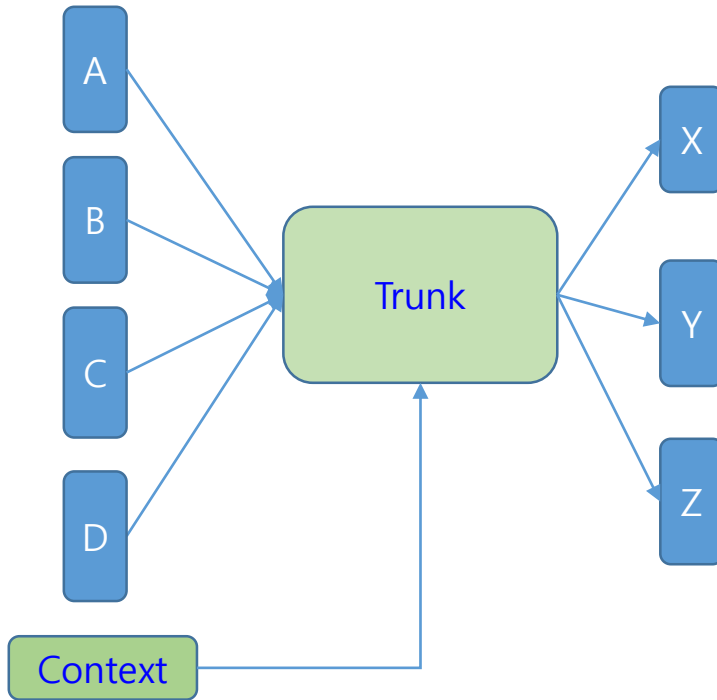
- Dynamic State: 신호 제어를 위한 교통 정보
- Static State: Context

개별 교차로의 특성: 평균 교통량, 용량 등

전이학습을 위한 최적화 모델(안)

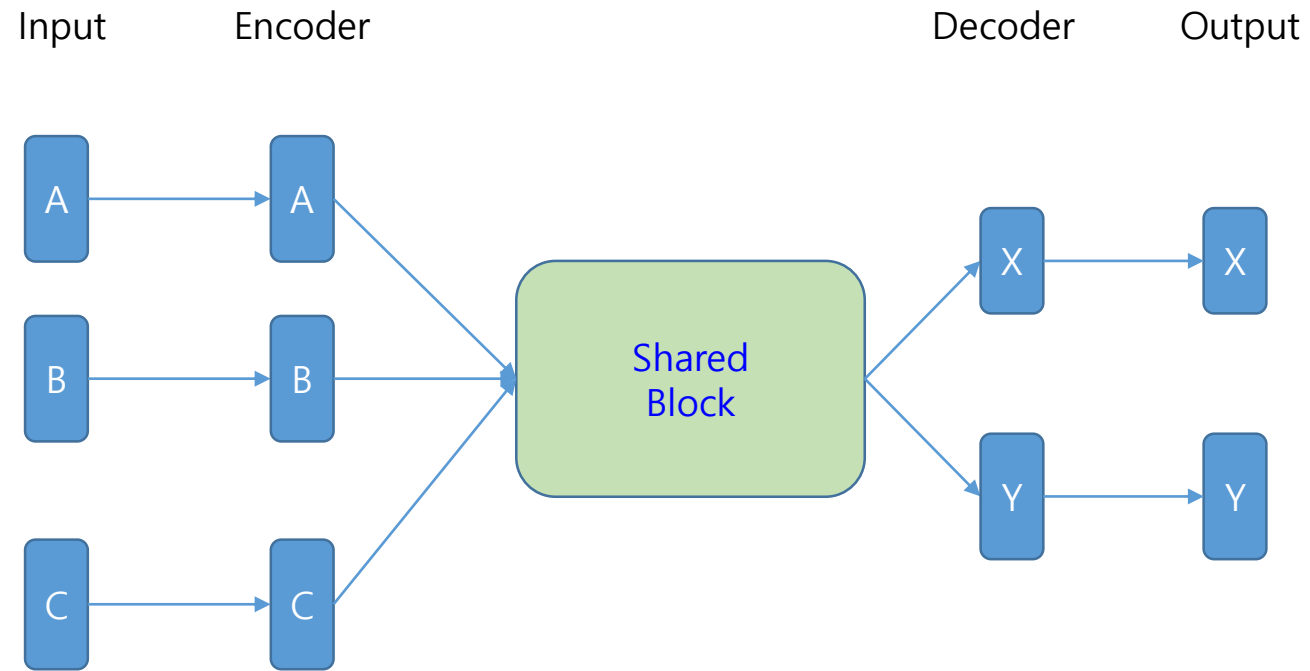
입력: State

출력: Action

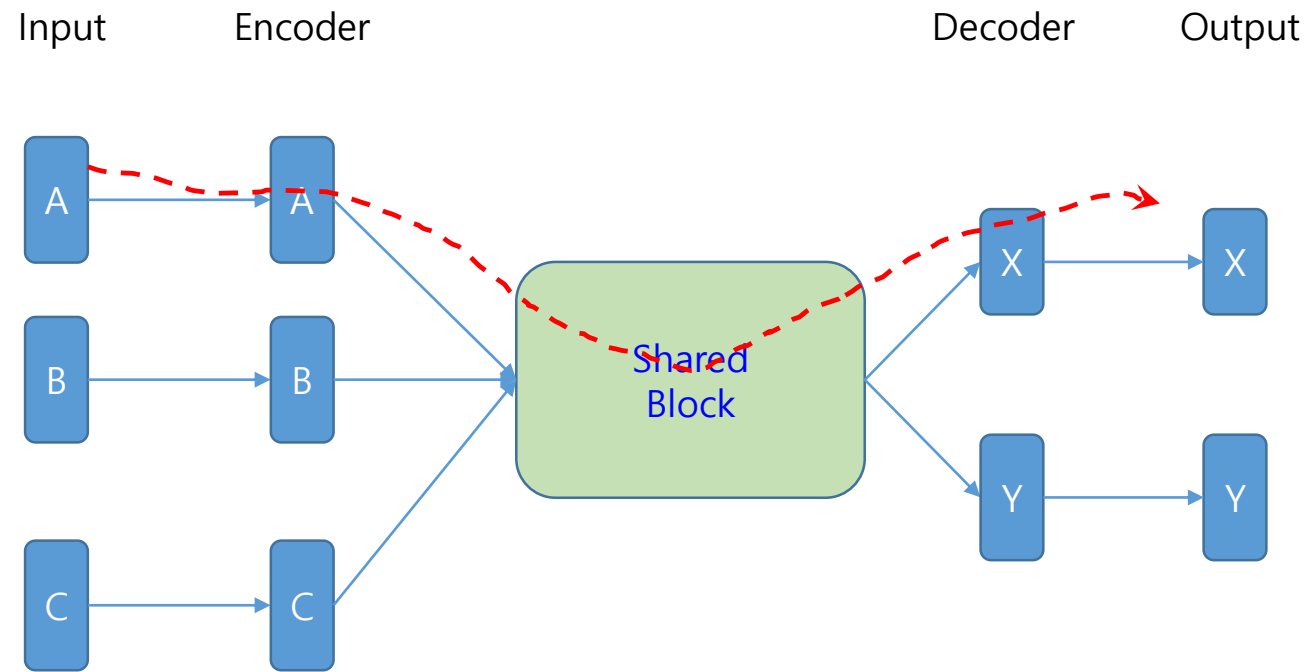


1. 교차로를 입력에 따라 A, B, C, D, ... type으로 구분
 - Type 별로 preprocessing/embedding block(layer) 구성
2. 교차로를 출력에 따라 X, Y, Z, ... type으로 구분
 - Type 별로 action 출력 block(layer) 구성
3. 교차로 type에 따라 입출력 block 을 선택하여 사용/학습
4. 'Context'와 'Trunk' block(layer)은 교차로 간에 공유 (shared block)
 - Target(Test) 교차로는 pre-trained 'Context'와 'Trunk' block을 재활용

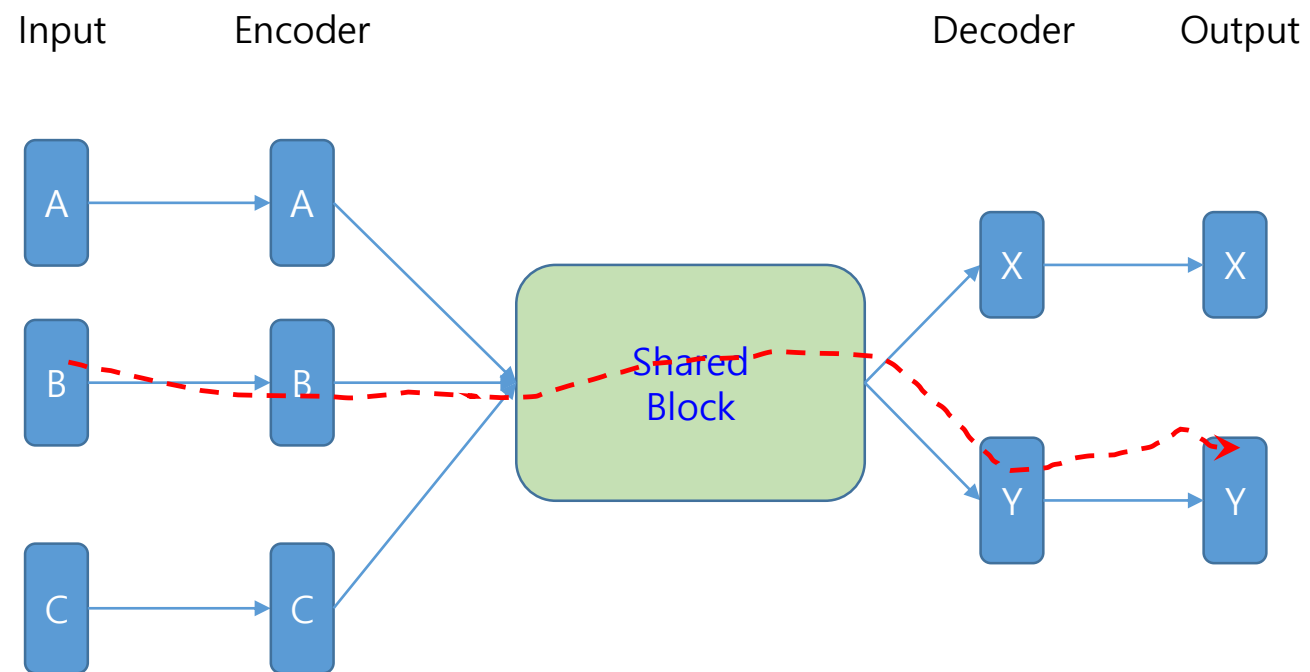
전이학습을 위한 최적화 DNN 구조



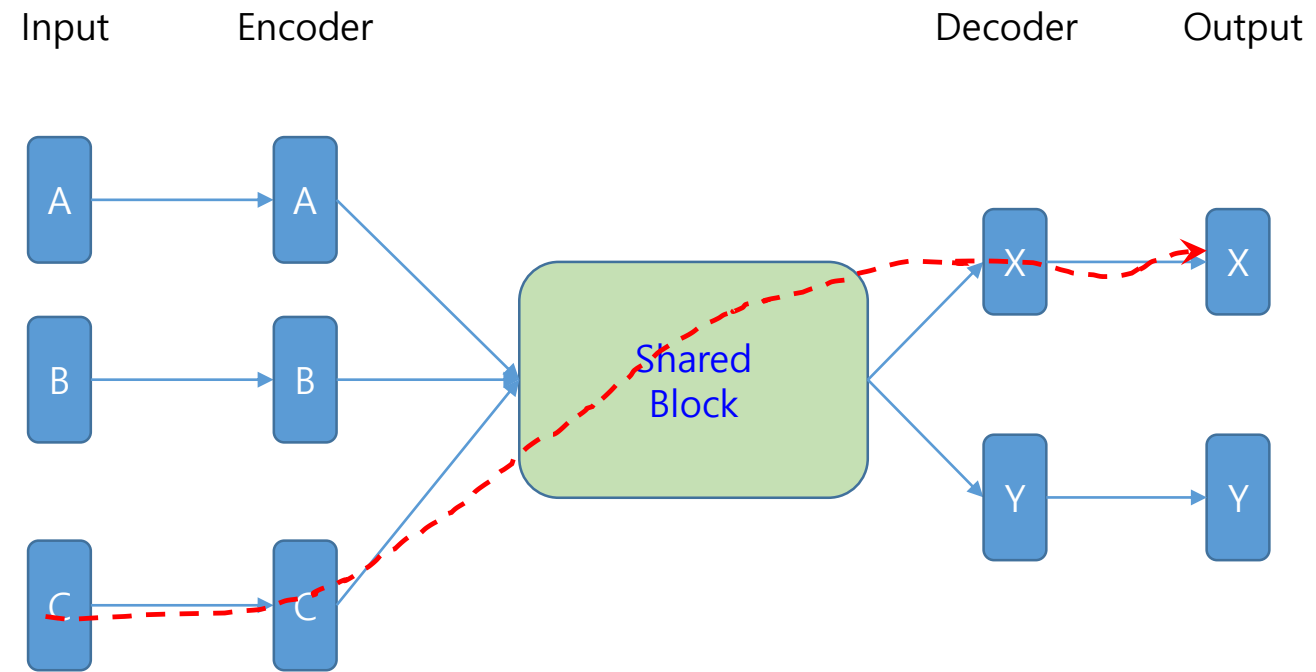
Input A, Output X



Input B, Output Y



Input C, Output X

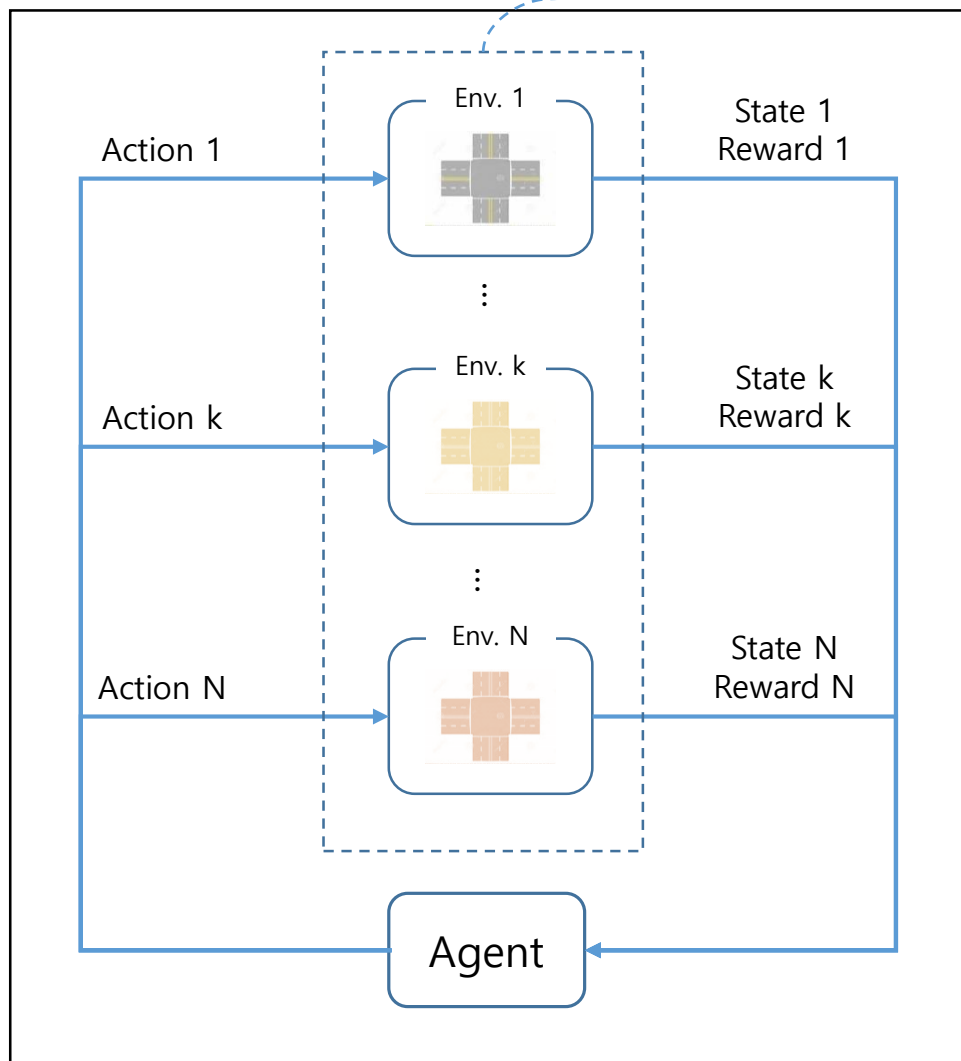


Q & A

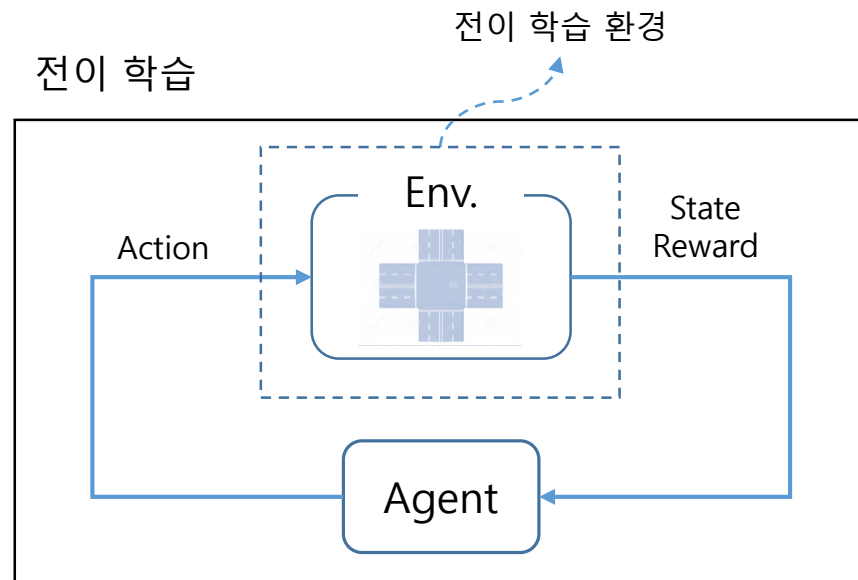
Transfer Learning

2023년 10월 25일

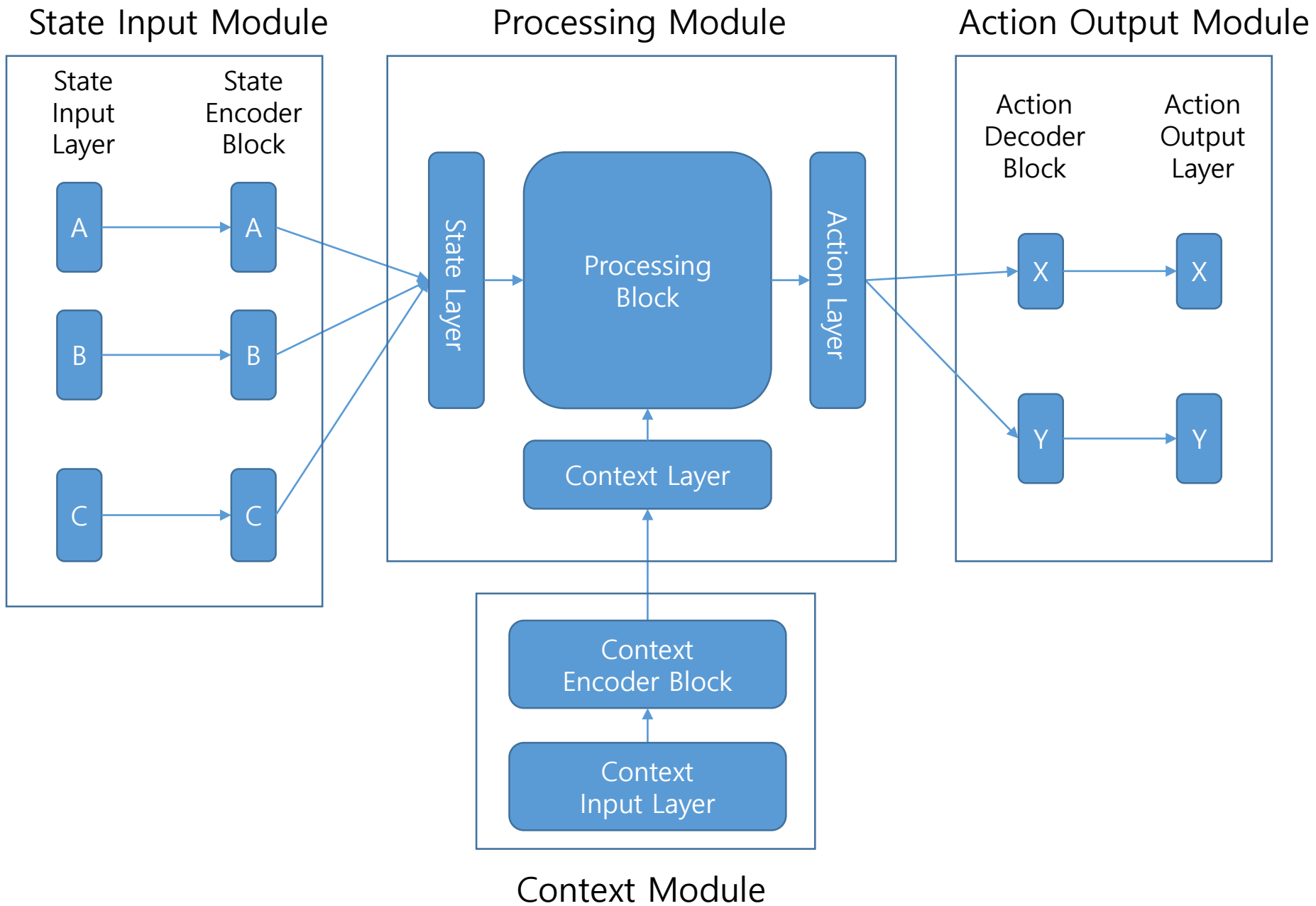
사전 학습

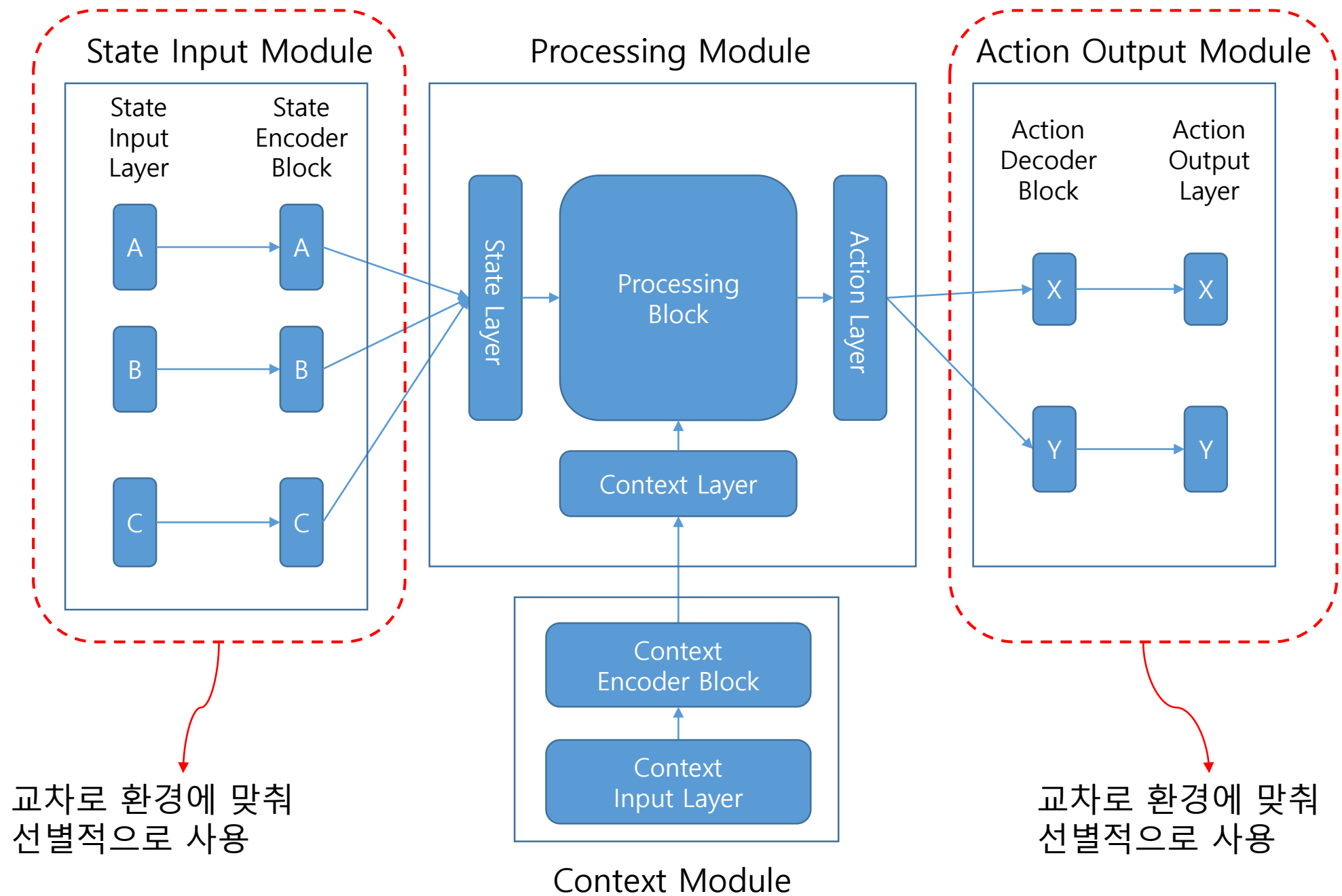


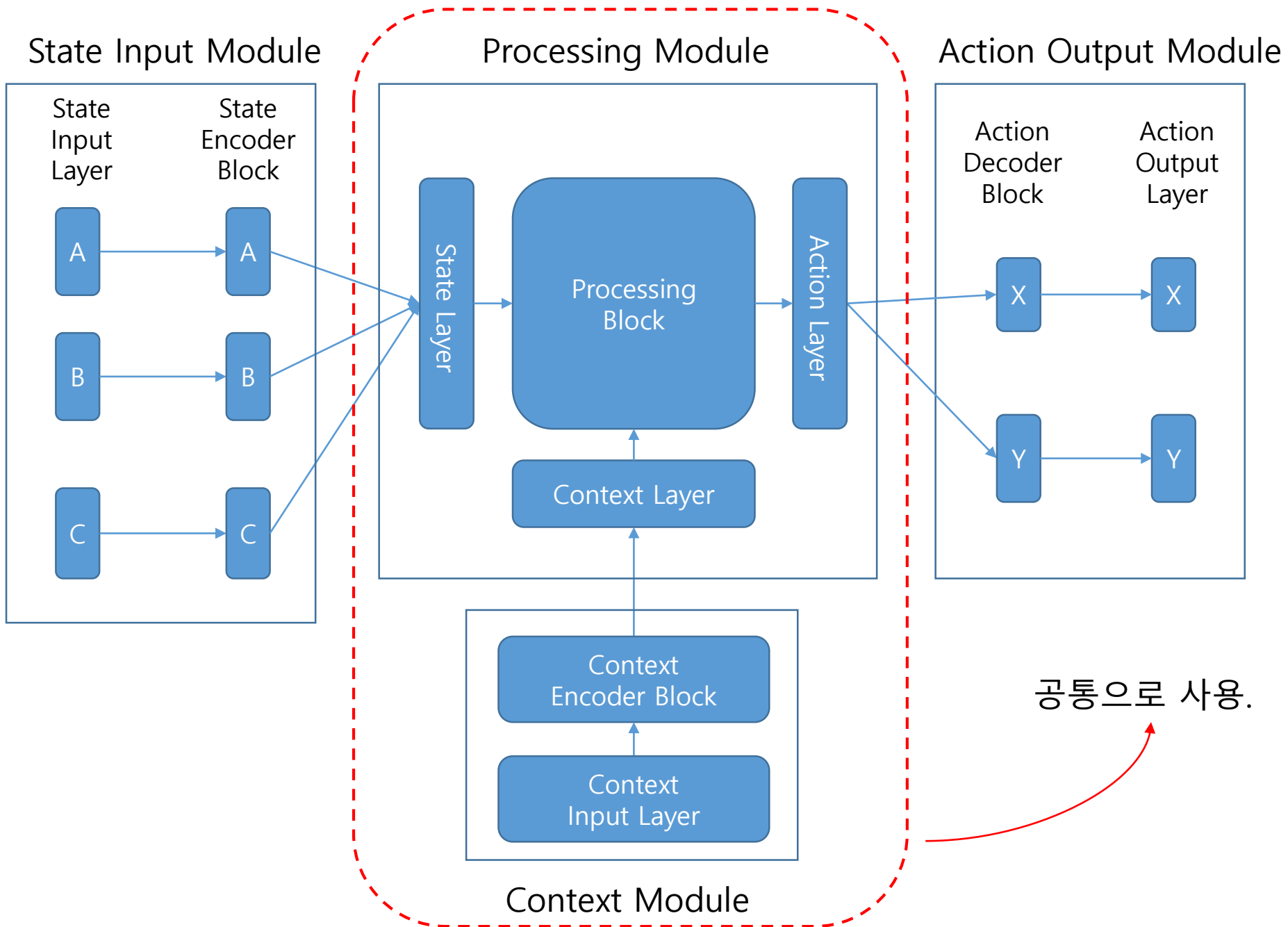
전이 학습



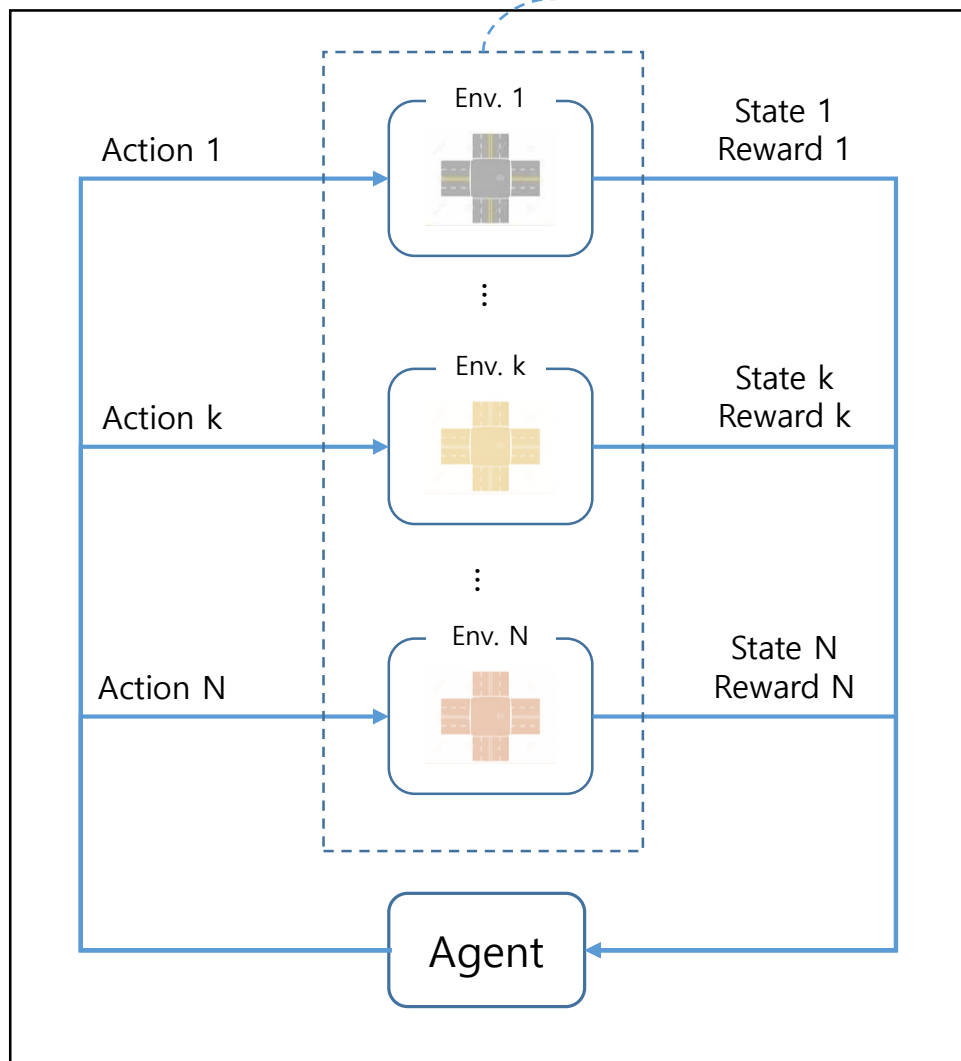
- 사전 학습
 - 다양한 교차로 환경에 대해서 학습
- 전이 학습
 - 새로운 교차로 환경에 적응 학습



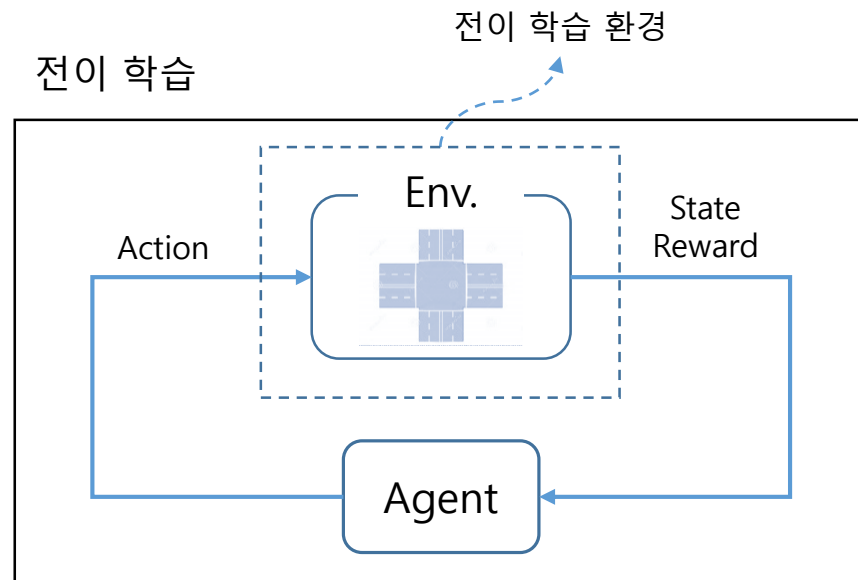




사전 학습



전이 학습



- 학습/실험 시간이 오래 걸림
 - 다양한 교차로 환경에서 학습
- 시뮬레이터-Agent 간 입출력 Interface 작업

추진 방향

- 전이 학습 적용 범위
 - 주기 또는 최소/최대 시간이 변경되면 처음부터 다시 학습해야 함.
 - 전이 학습을 적용하여 재학습 비용을 최소화.
- 사전/전이 학습 환경 구성
 - 교차로 고정.
 - 입출력(State, Action) 형식을 단일화
 - 기존 SAPPO Interface/DNN 모듈/학습 알고리즘 변경 최소화
 - 주기 또는 최대 시간을 다양하게 설정하여 다양한 환경 모사. 다른 특성(Context)은 고정.
 - 예) 현재 주기가 180초라면, 175, 177, 179, 181, 183, 185 등의 주기로 생성,
 - 예) 현재 교차로별 최대 시간에서 -3, -2, -1, +1, +2, +3
- 시뮬레이션 단위
 - 교차로 그룹(SA) 단위로 학습
 - 현재 실험 Code 변경 최소화
 - 기존 실험 결과 활용 가능. 실험 reference.

실험 설정

- 월평 SA28
 - Baseline. 기존 실험 결과 있음.
 - 성능이 좋음.
 - 비교적 학습 수렴 속도가 빠름
- 학습 환경 설정 1
 - 사전 학습 환경(학습): 현재 주기에서 -3, -2, -1, +1, +2, +3. 교차로 별 최소/최대는 고정.
 - 전이 학습 환경(테스트): 현재 주기(180초) 적용.
- 학습 환경 설정 2
 - 사전 학습 환경(학습): 현재 교차로 별 최대 시간에서 -3, -2, -1, +1, +2, +3. 최소 시간과 주기는 고정(180초)
 - 전이 학습 환경(테스트): 현재의 교차로별 최대 시간 적용.
- 필요 사항
 - TSS 파일 수정.

Q & A

신호 최적화 전이 학습 실험

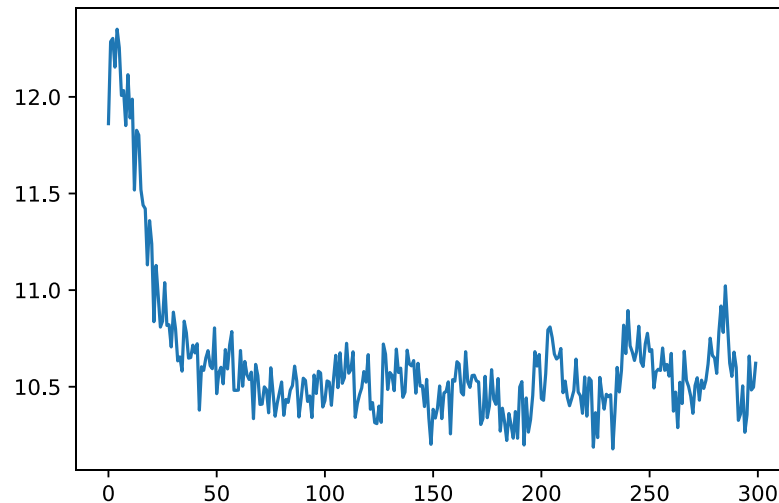
2023/11/22

주기 180초 학습 → 주기 185초 적용

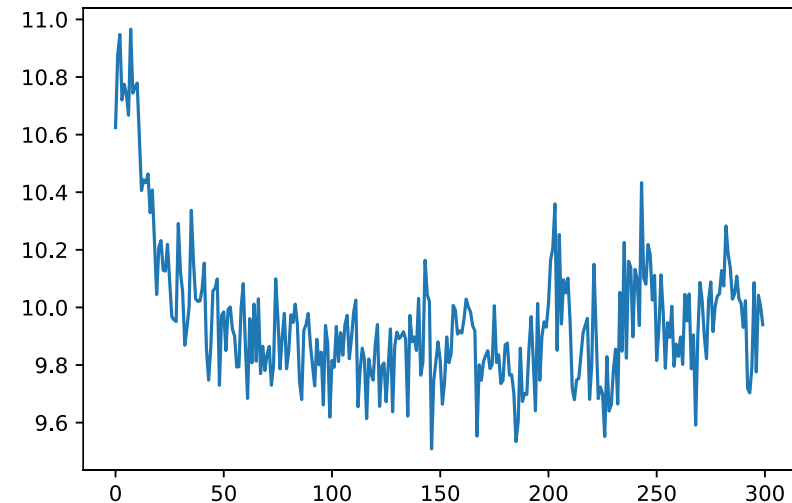
SA28. 주기 180초. 사전 학습 환경 및 사전 학습 모델

6월 수요, 10*146 epochs, Dirichlet(주기 고정), 주기 180초

수요	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.3167	10.2861	+22.76%	+18.30%
Test	12.6688	9.8848	+21.98%	+17.82% (Max. +19.04%)



Training

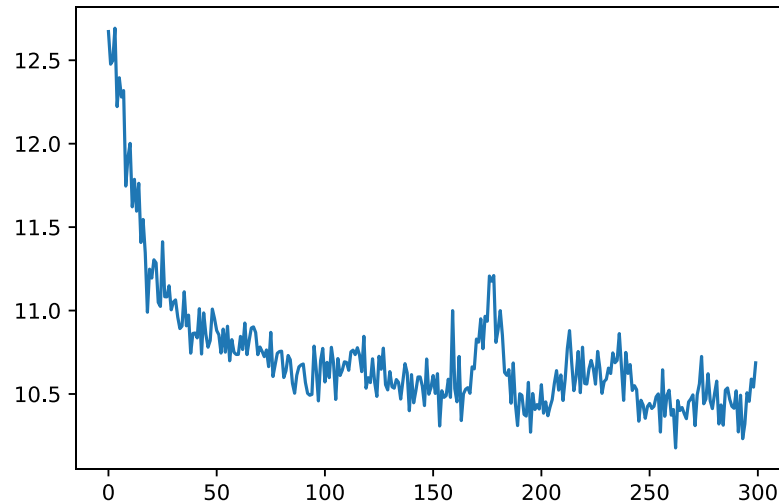


Validation

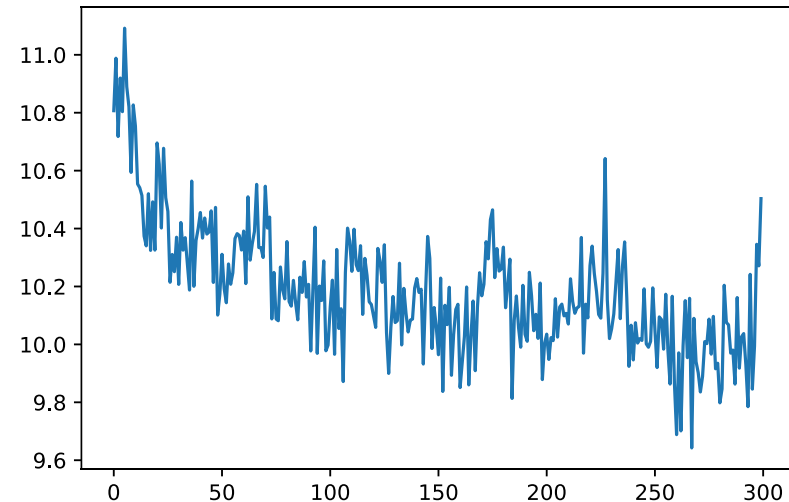
SA28. 주기 185초. 전이 학습 환경. Learning from scratch.

6월 수요일, 10*267 epochs, Dirichlet(주기 고정). 주기 185초

수요일	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.1753	10.2162	+22.46%	+17.71%
Test	12.5501	10.0025	+20.30%	+15.62% (Max. +18.69%)



Training



Validation

SA28. 주기 185초. 전이 학습 환경. Transfer Learning without Fine-tuning.

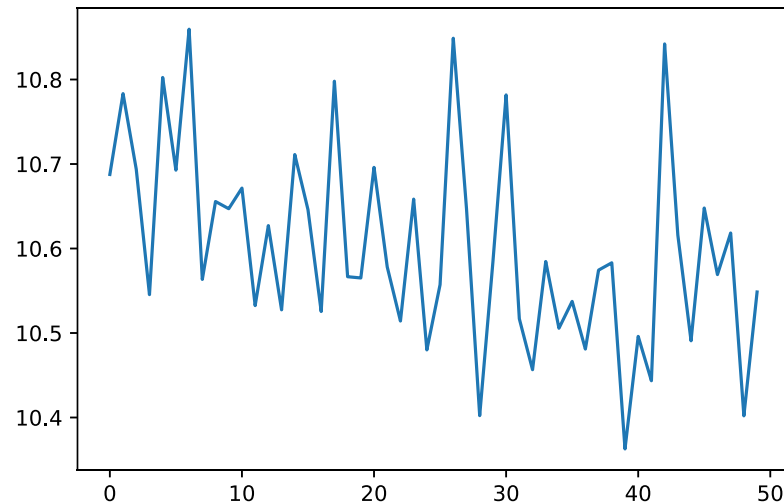
6월 수요, 주기 180초에서 모델 학습 -> 주기 185초에 적용

수요	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.1753	10.5278	+20.09%	+16.12%
Test	12.5501	10.2342	+18.45%	+14.41% (Max. +19.19%)

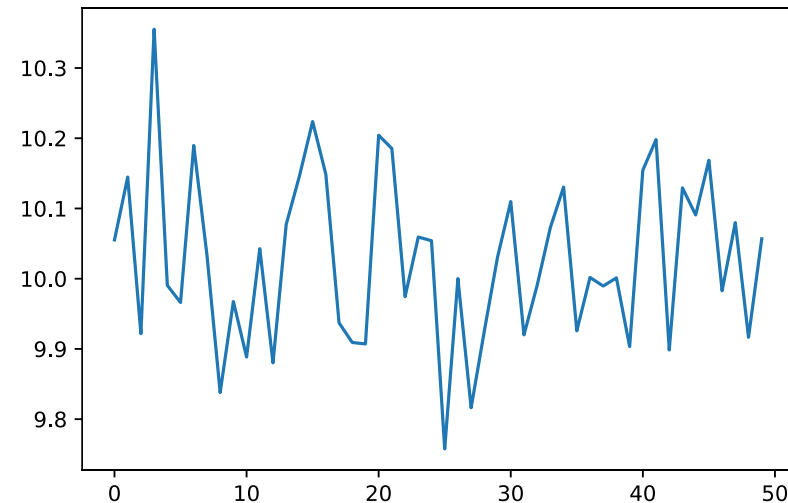
SA28. 주기 185초. 전이 학습 환경. Transfer Learning with Fine-tuning.

6월 수요일, 주기 180초에서 모델 학습 -> 주기 185초에 적용 (25*10 epochs)

수요일	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.1753	10.3398	+21.52%	+17.18%
Test	12.5501	10.0074	+20.26%	+15.85% (Max. +20.43%)



Training



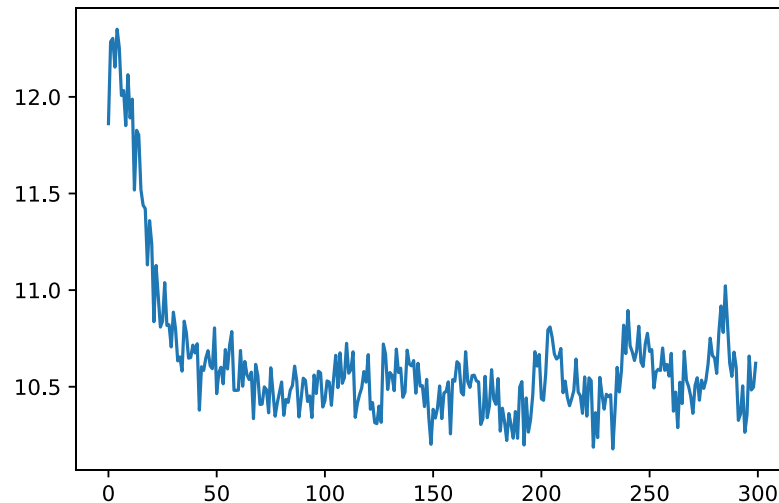
Validation

주기 180초 학습 → 주기 190초 적용

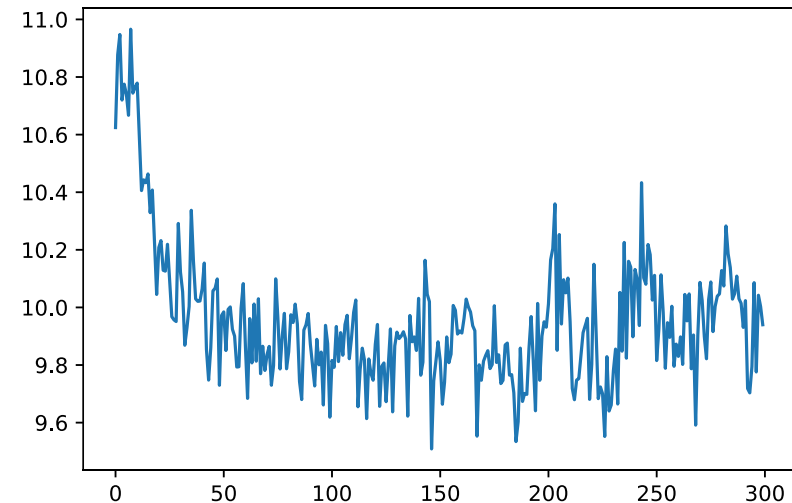
SA28. 주기 180초. 사전 학습 환경 및 사전 학습 모델

6월 수요, 10*146 epochs, Dirichlet(주기 고정), 주기 180초

수요	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.3167	10.2861	+22.76%	+18.30%
Test	12.6688	9.8848	+21.98%	+17.82% (Max. +19.04%)



Training

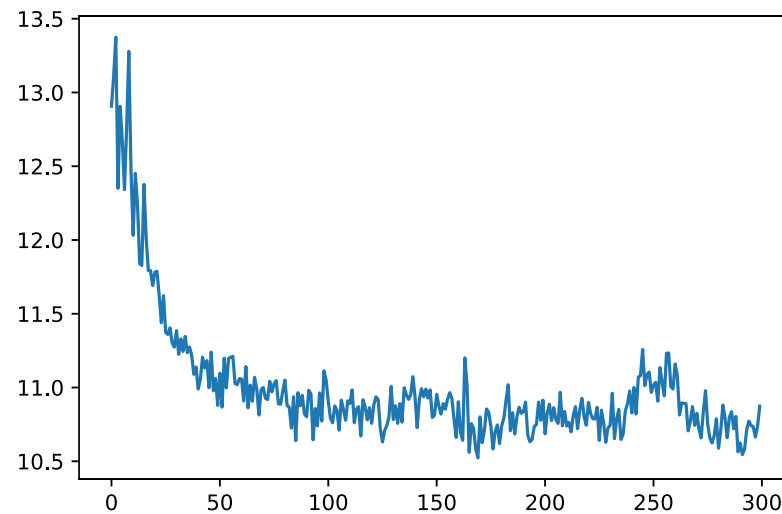


Validation

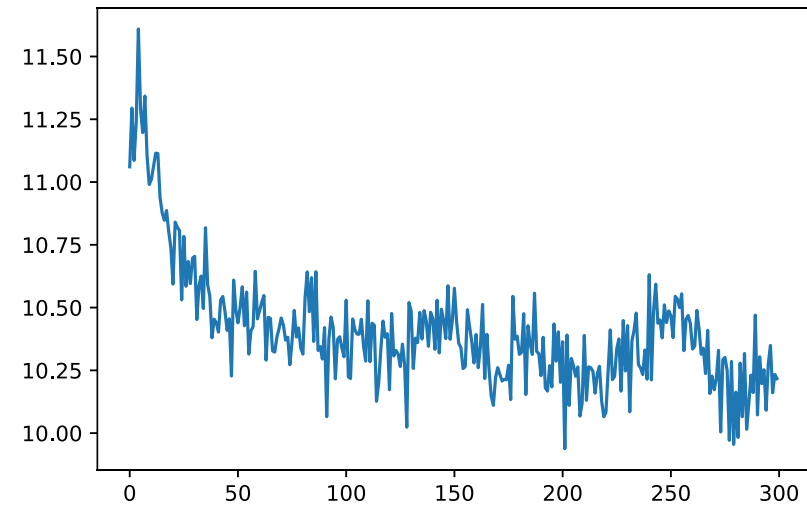
SA28. 주기 190초. 전이 학습 환경. Learning from scratch.

6월 수요일, 10*201 epochs, Dirichlet(주기 고정). 주기 190초

수요일	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.0199	10.3532	+20.48%	+16.57%
Test	12.6448	10.3705	+17.99%	+14.21% (Max. +19.31%)



Training



Validation

SA28. 주기 190초. 전이 학습 환경. Transfer Learning without Fine-tuning.

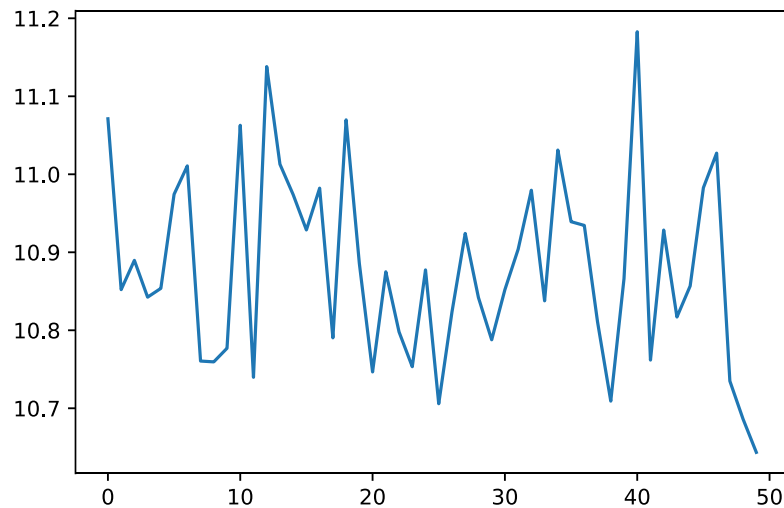
6월 수요, 주기 180초에서 모델 학습 -> 주기 190초에 적용

수요	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.0199	10.6720	+18.03%	+14.63%
Test	12.6448	10.6463	+15.80%	+12.87% (Max. +15.78%)

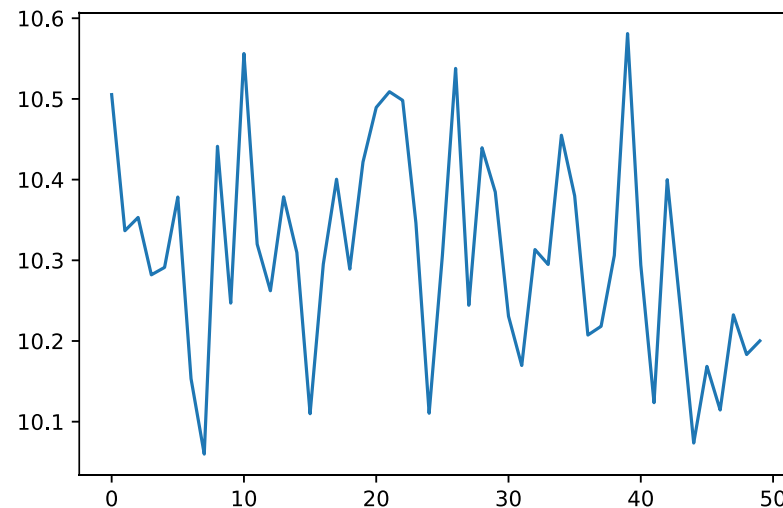
SA28. 주기 190초. 전이 학습 환경. Transfer Learning with Fine-tuning.

6월 수요, 주기 180초에서 모델 학습 -> 주기 190초에 적용 (7*10 epochs)

수요	Reward (WQ)		Improvement %	
	고정 신호	RL	Reward (WQ)	Avg. Travel Time
Training	13.0199	10.9252	+16.09%	+13.16%
Test	12.6448	10.3357	+18.26%	+14.88% (Max. +16.86%)



Training



Validation

Q & A