Lab seminar

RILS LAB

Robot Intelligence Learning & System

01

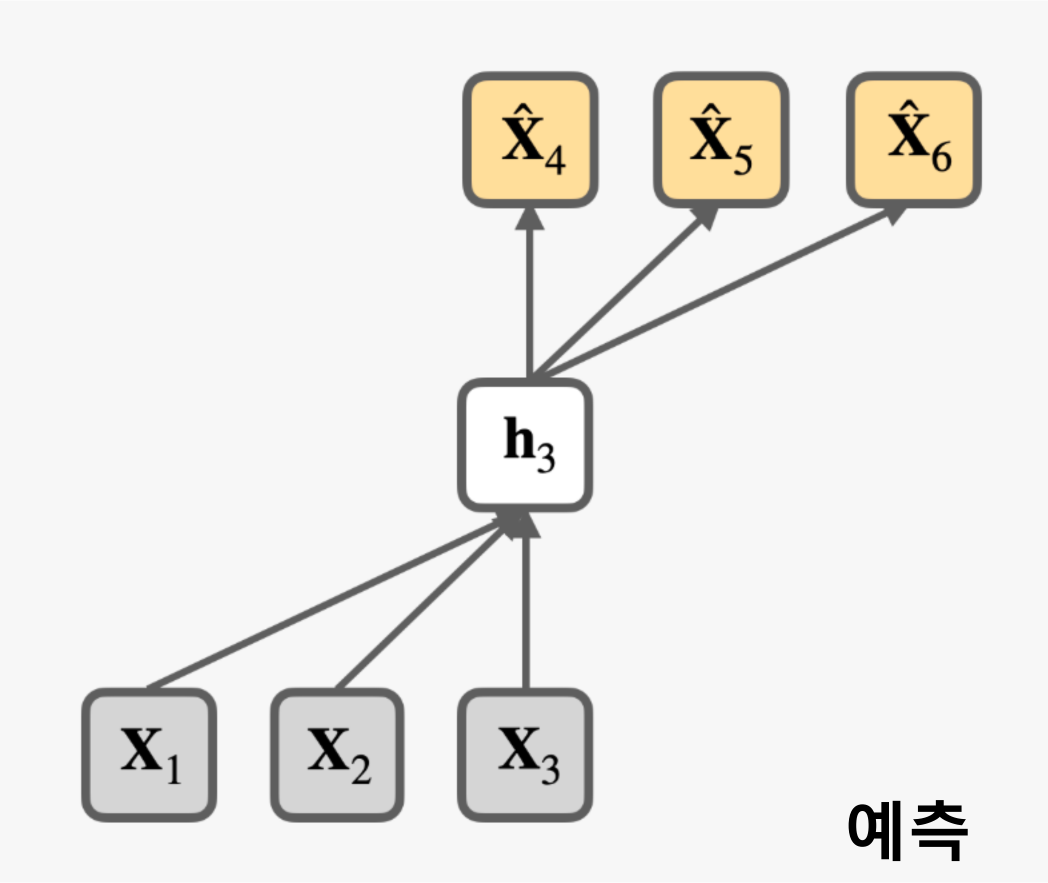
Learning to Theorize the World from Observation

Doojin Baek* Gyubin Lee* Junyeob Baek Hosung Lee Sungjin Ahn

ICML / 2026(version2)
https://arxiv.org/abs/2605.03413

논문 배경 :
Anti-Latent prediction world model

"이해 != 정확한 예측" in Developmental Cognitive Science

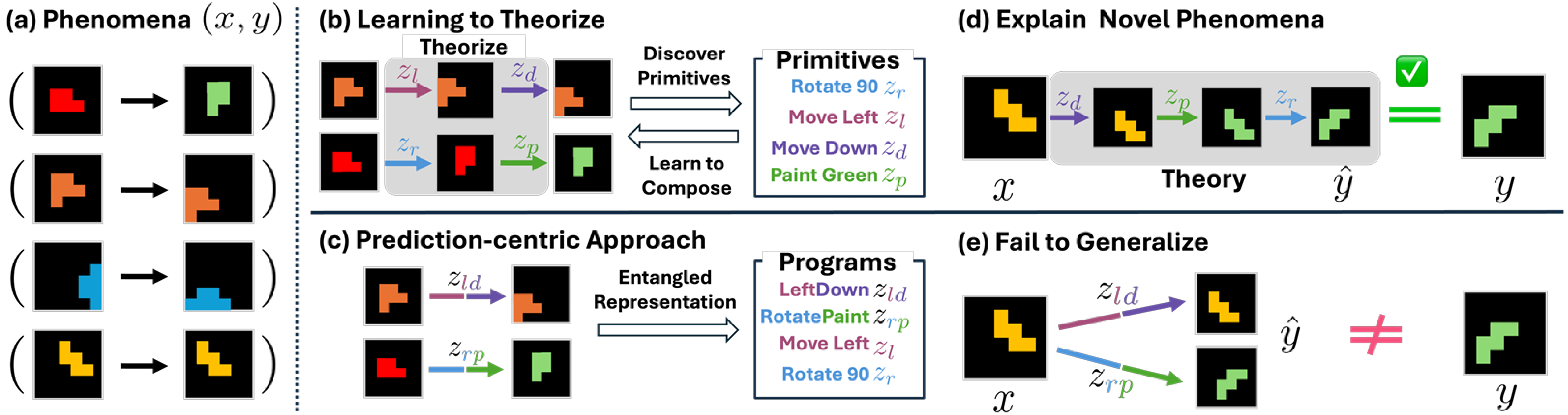


vs

Explain

논문 이론 :
순수 관측에서 시작되는 조합적 이론 구성

Explicit meaning dioxide



학습 과정 : 입력 → Encoder → Theorize → Decoder → MDL(가공) → Loss

논문 성능 : Generalization

α	Method	In-distribution		Comp. OOD		Length OOD	
		Self-Ex.	Transf.	Self-Ex.	Transf.	Self-Ex.	Transf.
0.33	Disc-Mono	**0.890**	0.668	0.005	0.004	0.012	0.009
	Cont-Mono	0.834	0.001	0.165	0.000	0.321	0.000
	Cont-Mono-Opt	0.866	0.001	0.218	0.000	0.394	0.000
	NEO (Ours)	0.847	0.792	0.357	0.345	0.045	0.038
	NEO-S (Ours)	0.857	**0.809**	**0.831**	**0.759**	**0.620**	**0.524**
0.66	Disc-Mono	0.737	0.572	0.005	0.002	0.006	0.004
	Cont-Mono	0.500	0.002	0.086	0.000	0.158	0.001
	Cont-Mono-Opt	0.565	0.002	0.122	0.000	0.216	0.001
	NEO (Ours)	0.794	0.731	0.609	0.573	0.023	0.019
	NEO-S (Ours)	**0.977**	**0.939**	**0.994**	**0.959**	**0.766**	**0.696**
1.00	Disc-Mono	0.662	0.475	·	·	0.006	0.004
	Cont-Mono	0.810	0.000	·	·	0.649	0.000
	Cont-Mono-Opt	0.846	0.000	·	·	0.743	0.000
	NEO (Ours)	0.724	0.675	·	·	0.025	0.023
	NEO-S (Ours)	**0.990**	**0.954**	·	·	**0.799**	**0.707**

02

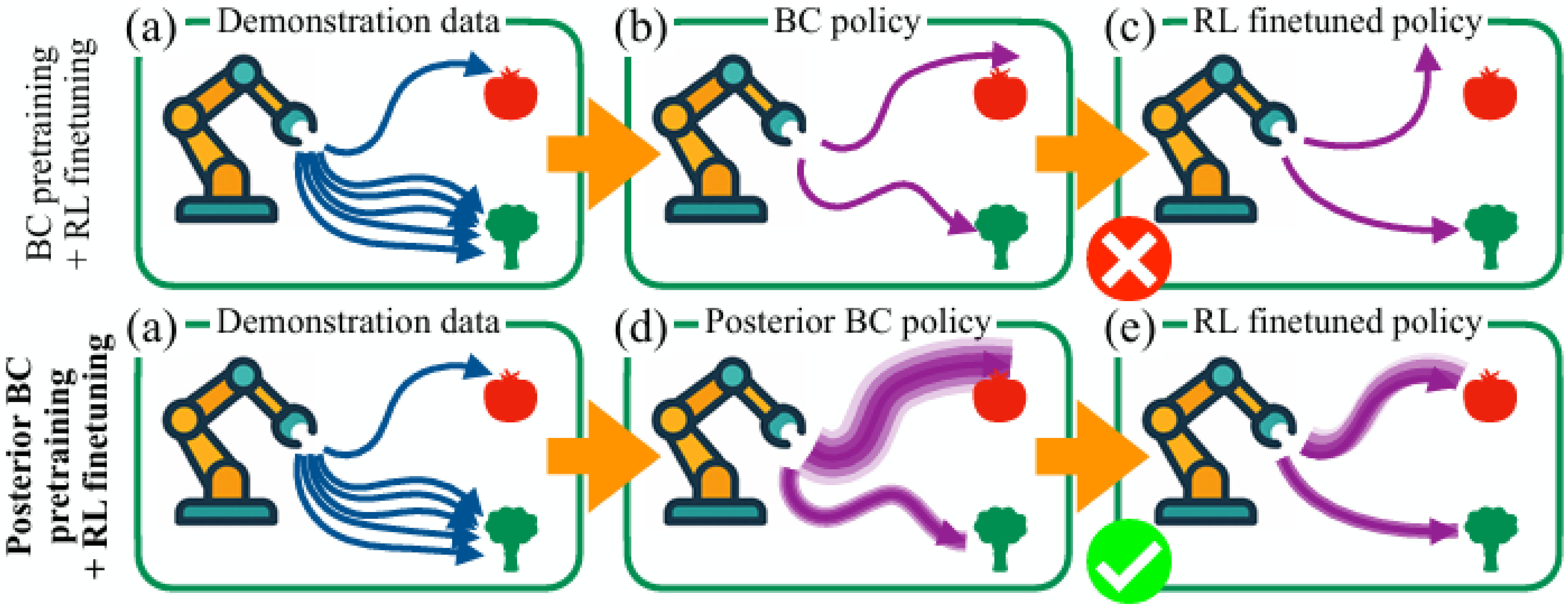
Posterior Behavioral Cloning: pretraining BC Policies for Efficient RL Finetuning

Andrew Wagenmaker* Perry Dong Raymond Tsao Chelsea Finn Sergey Levine

ICML / 2026
https://arxiv.org/pdf/2512.16911

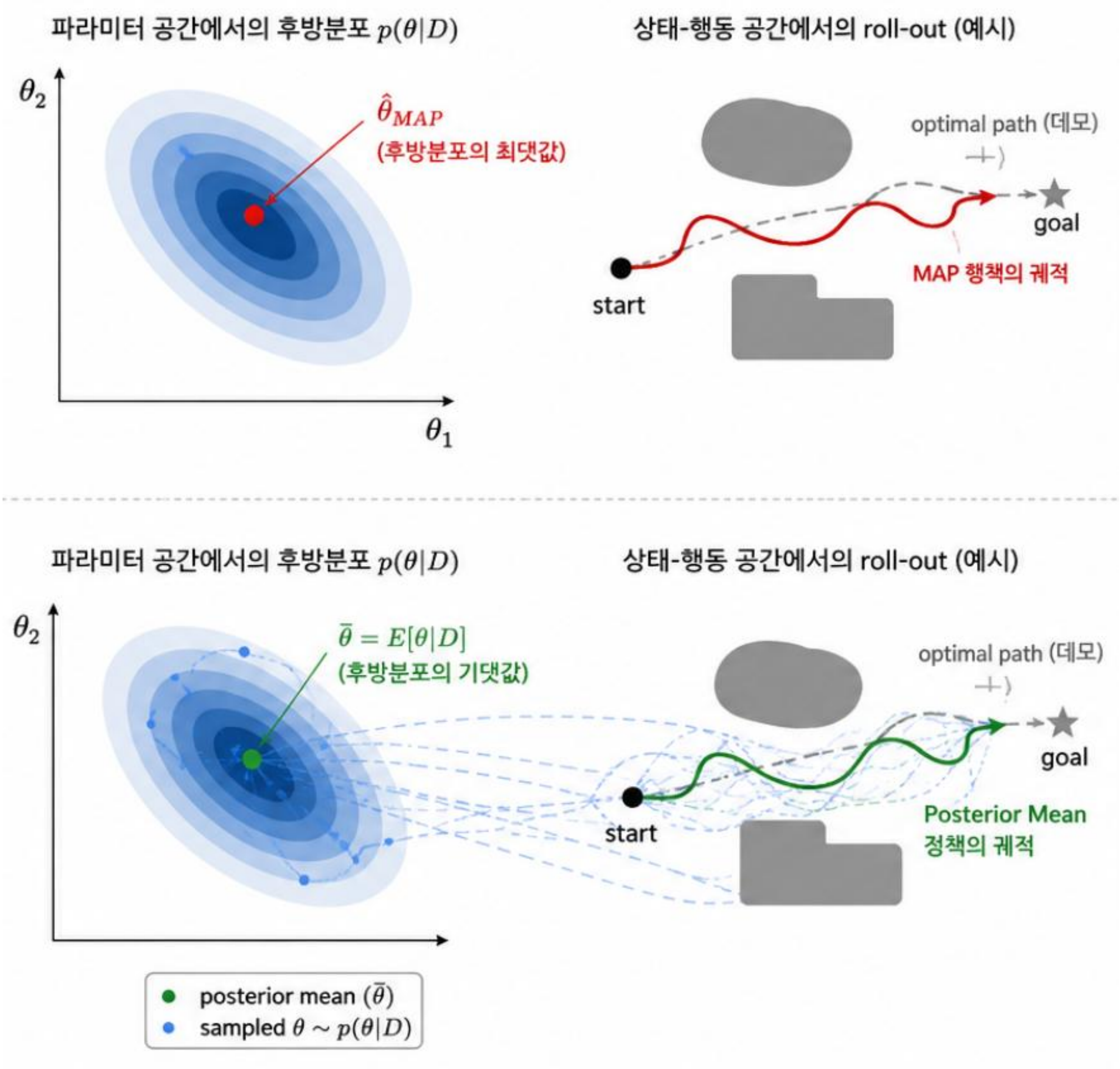
논문 배경 :
BC to Fine-tuning 시 문제점

Fine-tuning을 잘하려 해도 BC에서 망하면 의미가 없다



논문 이론 :
Post BC ➔ Fine-tuning

BC의 범위를 늘리자 : 적은 샘플은 불확실한 습관(posterior mean)을 나타낸다



학습 구조 : 경량모델로 사전 MAP 추출
→ Ensemble 기반 posterior mean 추출
→ 실제 학습시 target에 noise

03

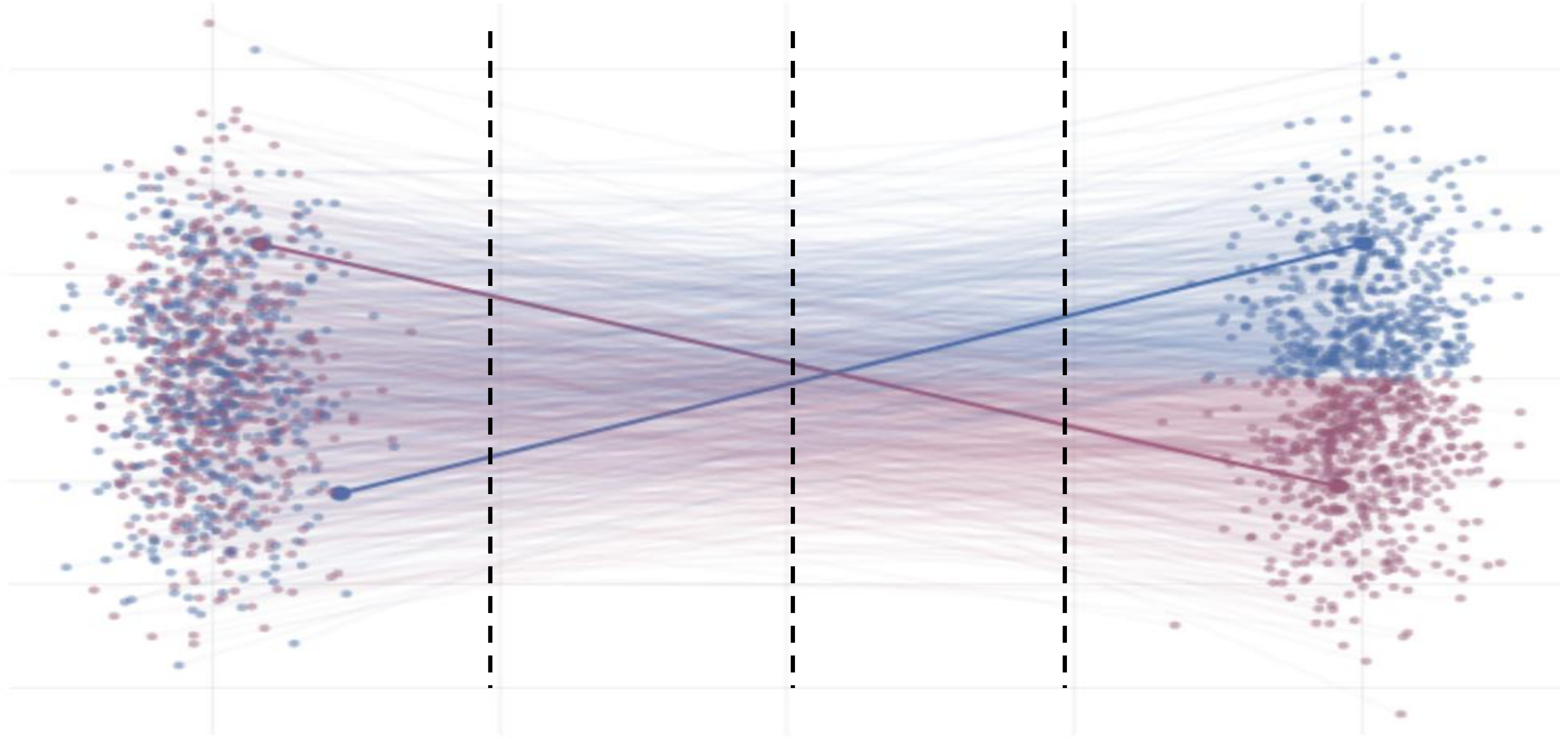
Q-learning with adjoint matching

Qiyang Li Sergey Levine

ICLR / 2026
https://arxiv.org/pdf/2601.14234

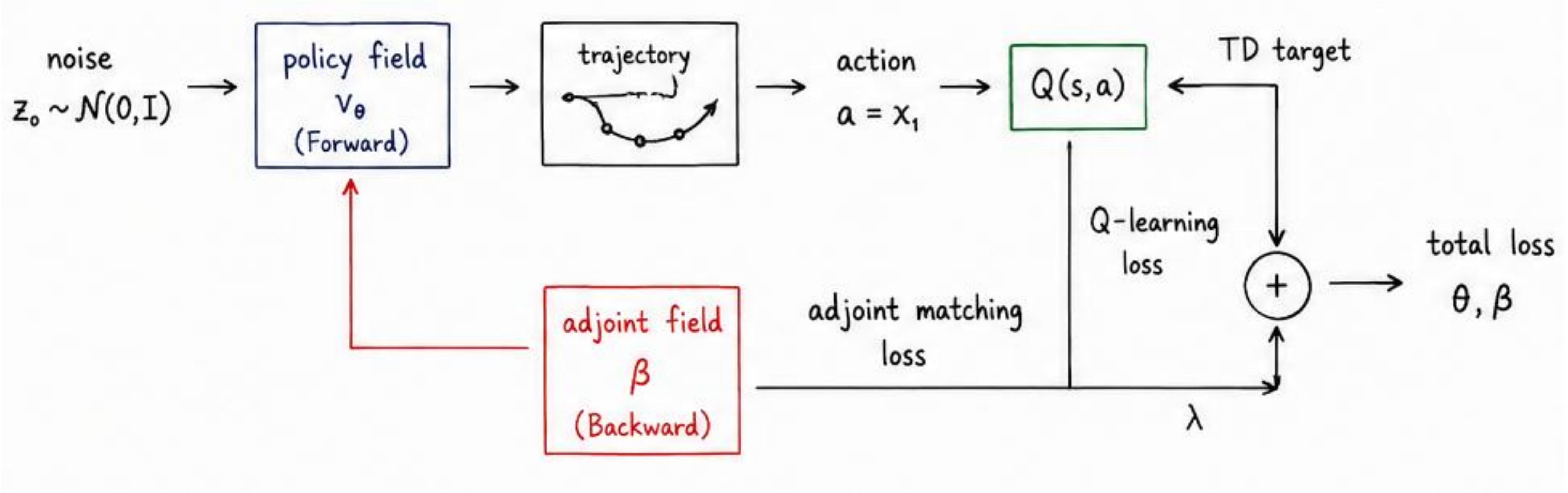
논문 배경:
Flow matching Policy과 Q-learning의 결합 불안정성

ODE trajectory에서의 BPTT로 인한 발산



논문 구현:
Backprop을 대신할 별도의 vector field 학습

대체 신경망은 자기 참조를 고정시켜 발산 억제



04

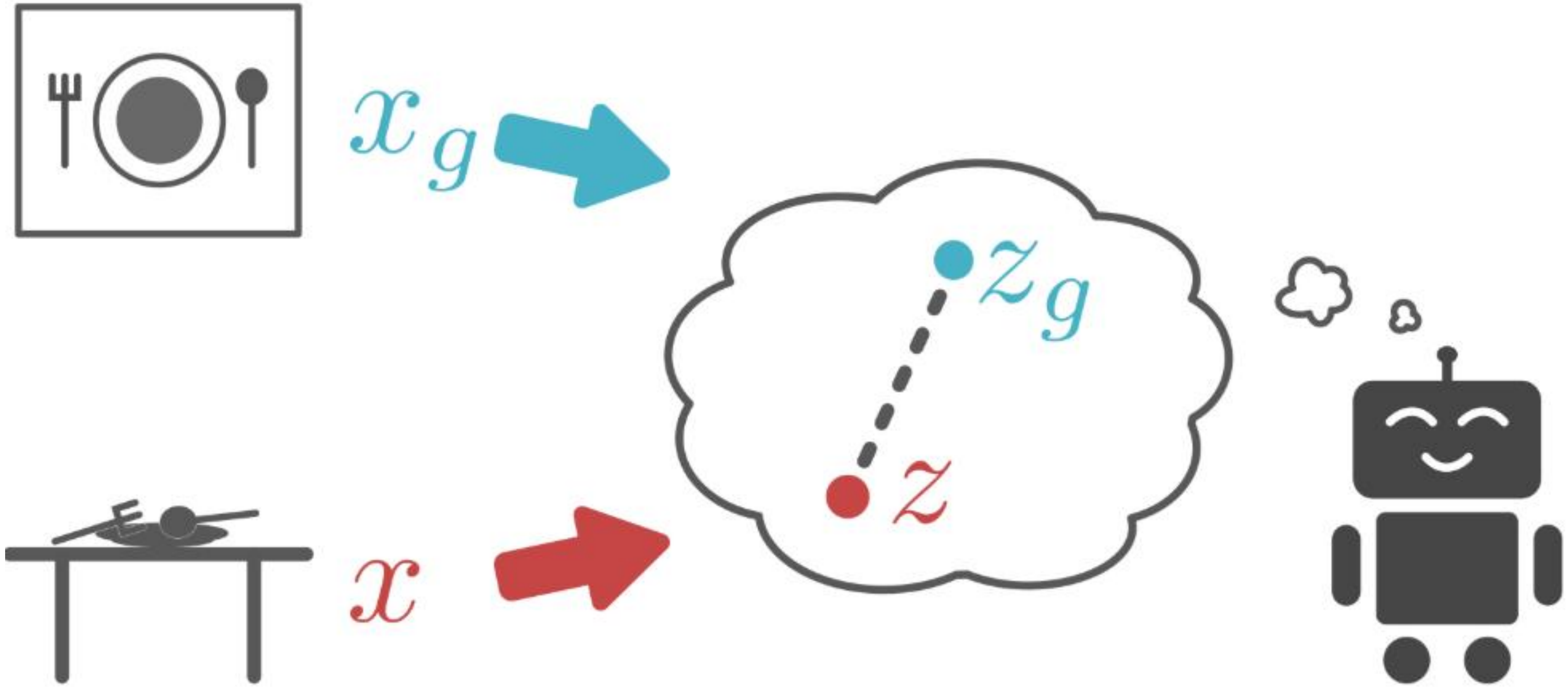
Scaling Goal-conditioned RL with Multistep Quasimetric Distances

Bill Chunyuan Zheng Vivek Myers Benjamin Eysenbach Sergey Levine

ICLR / 2026
https://openreview.net/pdf?id=UElh7vzgKX

논문 배경 :
Long horizon problem in Goal-conditioned RL

Latent 공간 학습시 일반적인 유클리디안 거리로는 현실적 제약 표현이 어렵다



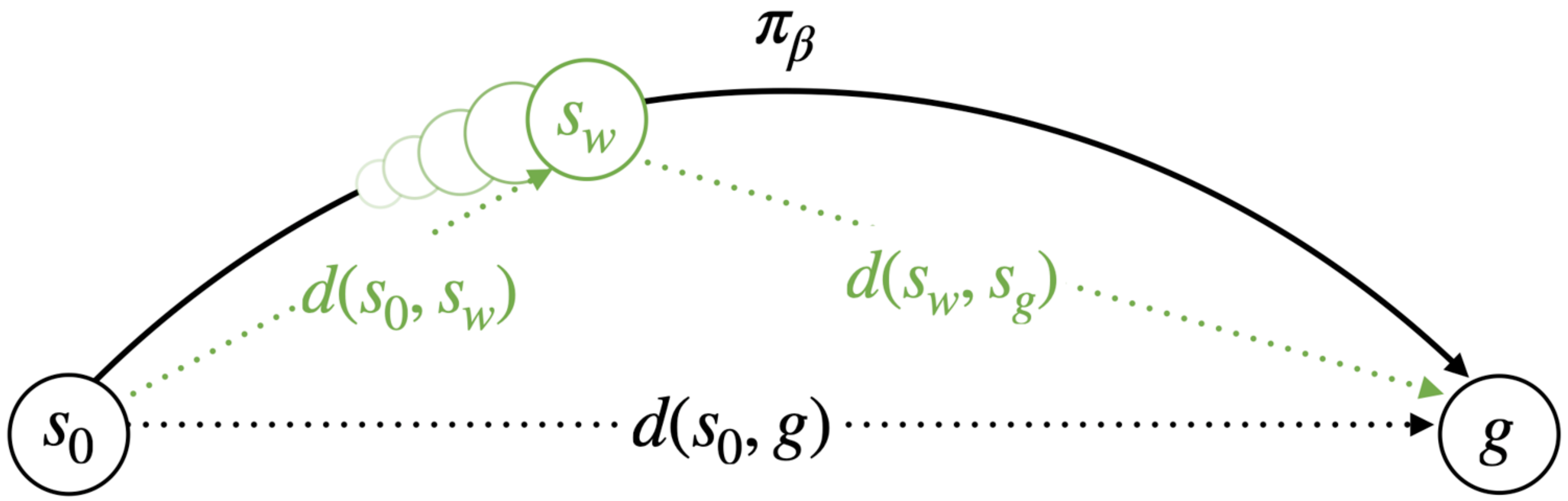
입력

encoder

think

논문 구현:
Quasimetric + Multistep Structure

“구조화된 인코더 + 확률적인 목표"



Quasimetric Architecture: $d(s_0, g) \leq d(s_0, s_w) + d(s_w, s_g)$

$$+ \quad \mathrm{Geom}(1-\gamma)$$

추가점 : batch(N) → N^2의 연산과 LINEX 연산

05

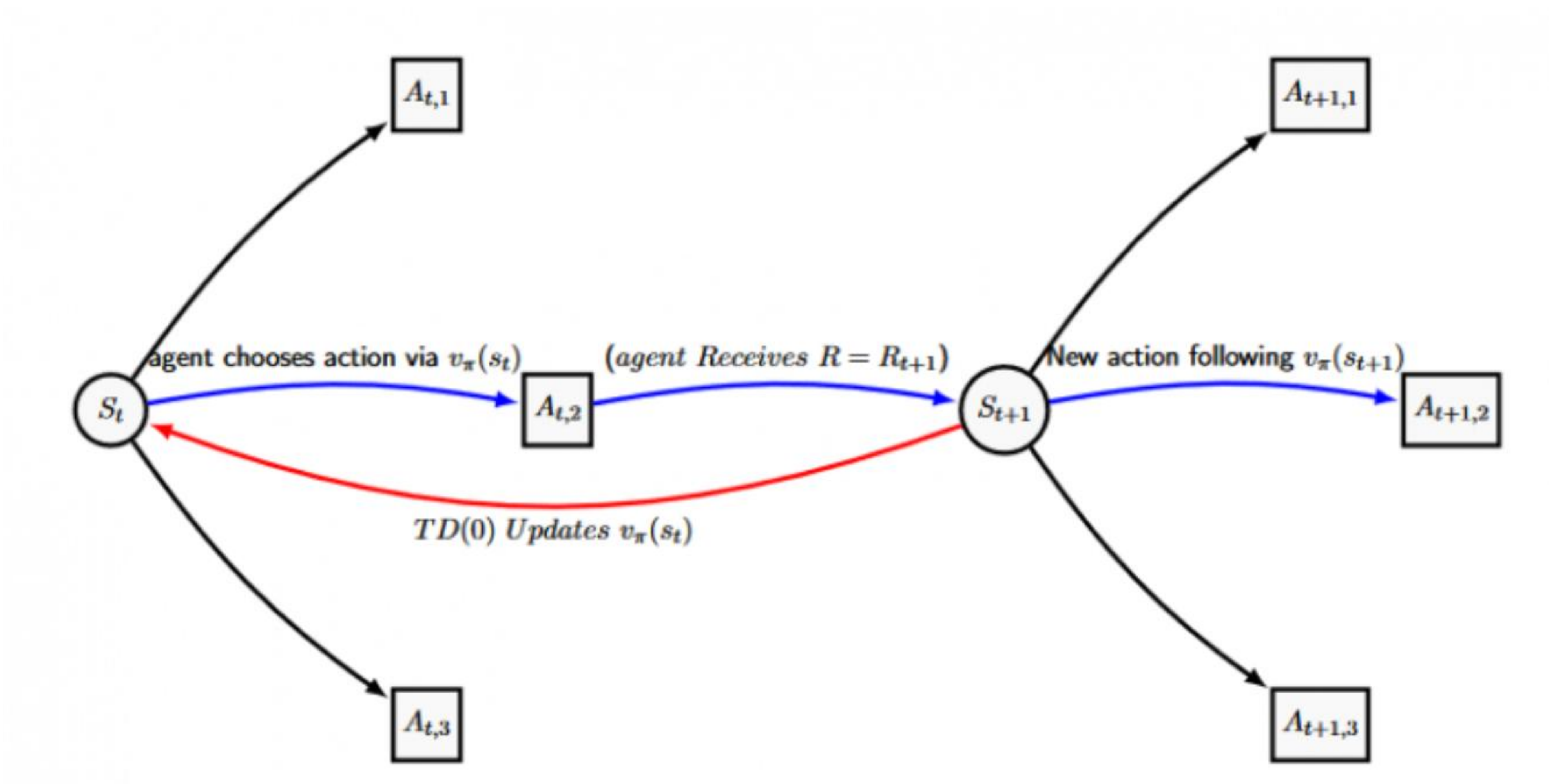
Temporal Difference Flows

JesseFarebrother, MatteoPirotta, AndreaTirinzoni, RémiMunos, AlessandroLazaric,
Ahmed Touati

ICML / 2026
https://arxiv.org/pdf/2503.09817

논문 배경:
Long horizon problem in TD structure

one-step prediction으로 장기 미래를 예측하는 것의 한계점

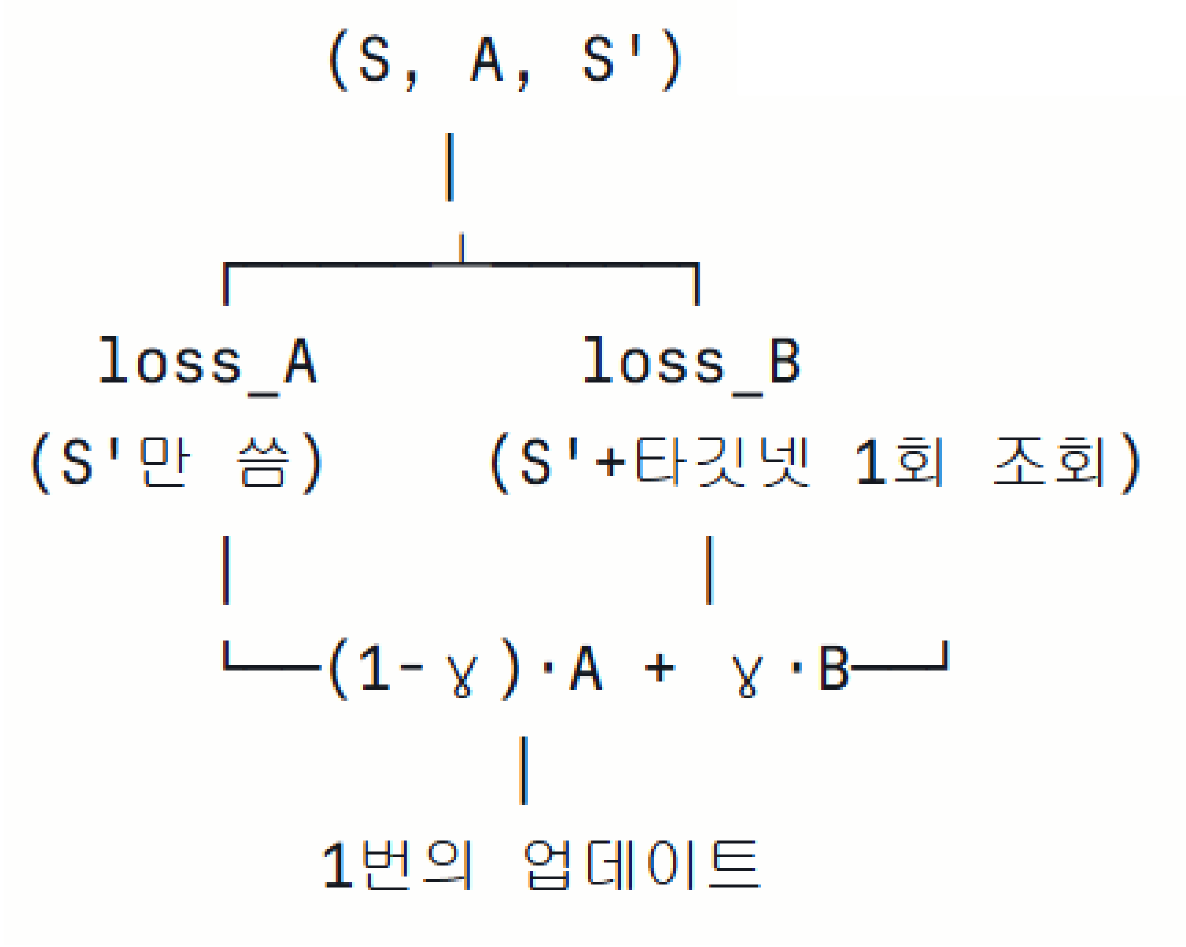


논문 이론:
GHM with flow matching (extra model)

"GHM : Geometric Horizon Measures" ➔ 미래 상태 분포 모델링

방법 :
확실한 1-step 미래와
불확실한 k-step 미래 속도 벡터를 섞어서 target

업데이트 : 방향 조정

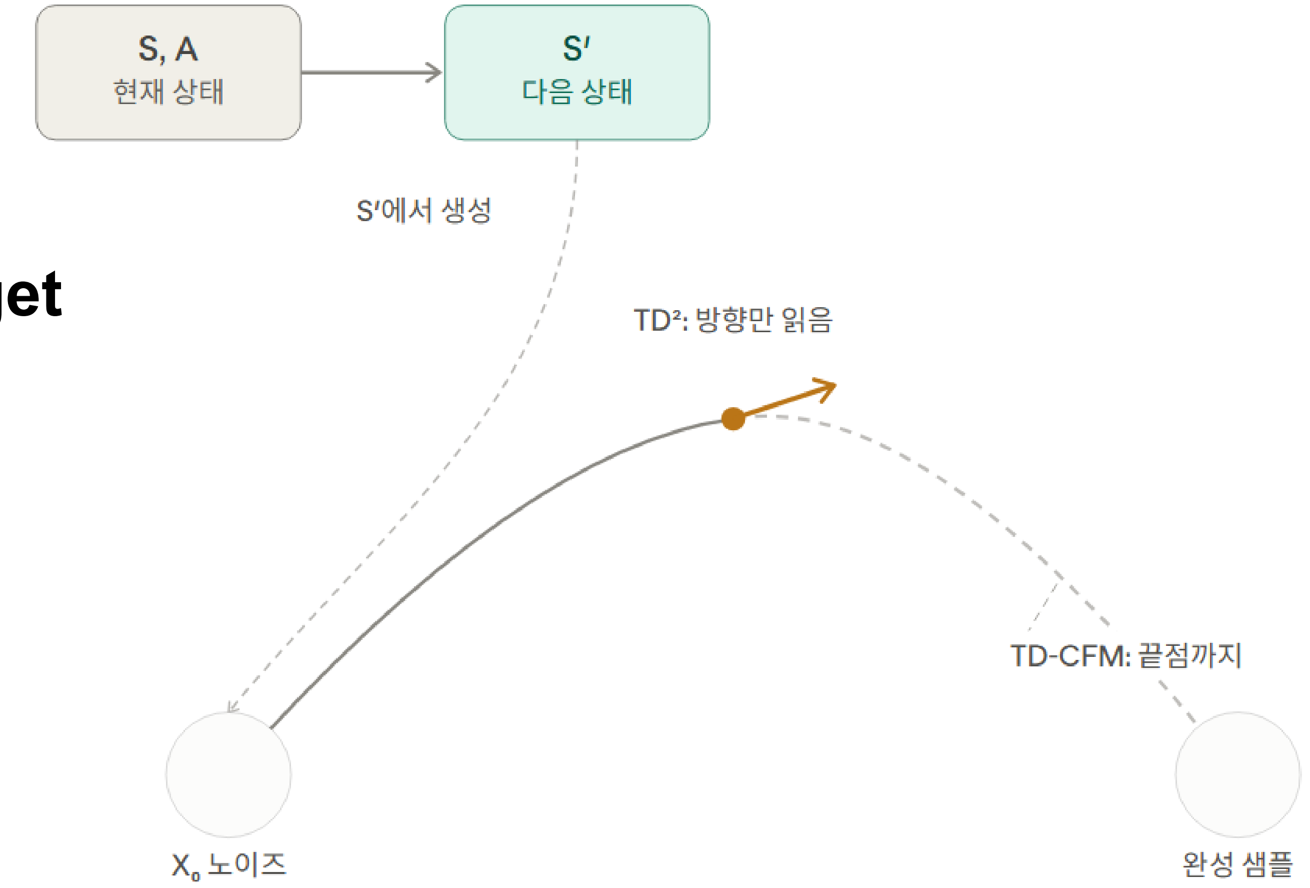


논문 이론:
GHM with flow matching (extra model)

“GHM : Geometric Horizon Measures“ ➔ 미래 상태 분포 모델링

방법 :
확실한 1-step 미래와
불확실한 k-step 미래 속도 벡터를 섞어서 target

업데이트 : 방향 조정



Thank you

Question & Disscusion

RILS LAB