R2P2: A ReparameterRized Pushforward Policy for Diverse, Precise Generative Path Forecasting

Dataset: http://www.nec-labs.com/~mas/R2P2
Project: http://www.cs.cmu.edu/~nrhineha/R2P2.html

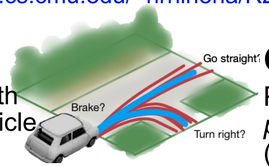
Nicholas Rhinehart[1,2], Kris M. Kitani[1], Paul Vernaza[2]. [1]Carnegie Mellon University, [2]N.E.C. Labs America
{nrhineha,kkitani}@cs.cmu.edu, pvernaza@nec-labs.com

Presented at ECCV 2018, Munich, Germany

Carnegie Mellon University
The Robotics Institute
NEC Laboratories America
Relentless passion for innovation

Problem:

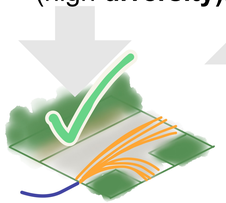
forecast a future path distribution of a vehicle given visual input



$$p(x|\phi), x = [x_1, \ldots, x_T], x_t \in \mathbb{R}^2$$

Goals:

Produce all *plausible* paths (high **diversity**)...



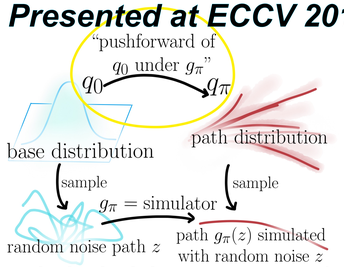
... while not producing any *implausible* paths (high **precision**).

Key ideas: learn a **conditional generative distribution** with an **invertible pushforward model.** Symmetric cross-entropy objective encourages **diversity** and **precision**.

$$\min_\theta H(p, q_\theta) + \beta H(q_\theta, \tilde{p})$$

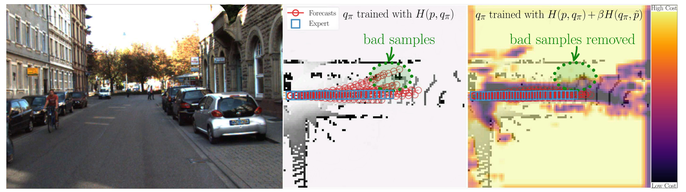
encourages diversity/coverage of training data

encourages precision: prevents "bad" samples



Key technical trick: represent model PDF q as the *pushforward* of a simple distribution under an invertible warp (simulator), enabling efficient evaluation of q and differentiation using the reparameterization trick.

Findings: Our models are best under both **distribution metrics** and **sample metrics,** outperforming state-of-the-art methods. Also, incorporating reverse cross-entropy improves samples.



Objective and Model

We represent our model, q, as the distribution of roll-outs of a **policy** π to be learned. This distribution is known as the "pushforward" of the base distribution, $z \sim q_0 = \mathcal{N}(0, I)$ under the **invertible roll-out simulator**,

$$x = g(z|\phi; \theta), z = g^{-1}(x|\phi; \theta)$$

A policy maps past states and context (causal information) to a next-state distribution

$$\phi = (x_{\text{past}}, M) \quad \psi_t = (x_{1:t-1}, \phi)$$

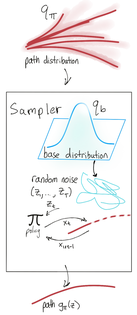
context — causal information

$$x_t = \mu_t(\psi_t; \theta) + \sigma_t(\psi_t; \theta) z_t$$

one-step policy

$$x_t \sim \mathcal{N}(\mu(\psi_t; \theta), \sigma_t(\psi_t; \theta)\sigma_t(\psi_t; \theta)^T)$$

next-state conditional distribution



Learning $\min_\theta H(p, q_\theta) + \beta H(q_\theta, \tilde{p})$

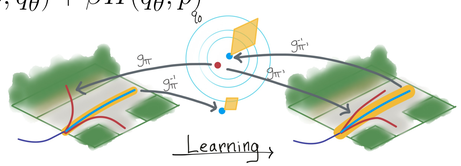
How can we optimize this objective with stochastic gradient descent? It involves the PDF of the trajectory distribution (**which most generative models can't compute**), and a complicated gradient of an expectation under a sampling distribution.

We can easily compute an unbiased estimate of $H(p, q_\theta)$

$$\log q(x|\phi) = \sum_{t=1}^{T} \log \mathcal{N}(\mu(\psi_t; \theta), \sigma(\psi_t; \theta)\sigma(\psi_t; \theta)^T)$$

Or, we use the invertibility of the roll-out

$$\log q(x|\phi) = \log \frac{q_0(g^{-1}(x|\phi; \theta))}{|\det J_g(g^{-1}(x|\phi; \theta))|}$$



Key idea: the "roll-out" view allows us to apply the **reparameterization trick** to obtain a low-variance estimate of $\nabla_\theta H(q_\theta, p) = -\nabla_\theta \mathbb{E}_{x \sim q_\theta} \log \tilde{p}(x)$

these don't commute — *these do commute*

$$\boxed{\nabla_\theta \mathbb{E}_{x \sim q_\theta}} \log \tilde{p}(x) = \boxed{\nabla_\theta \mathbb{E}_{z \sim q_0}} \log \tilde{p}(g(x|\phi; \theta))$$

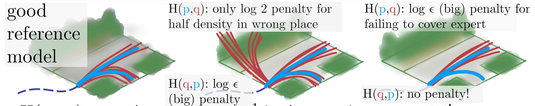
Cross-Entropies Properties

forward cross-entropy: $H(p, q_\theta) = -\mathbb{E}_{x \sim p} \log q_\theta(x)$

-sensitive to distribution diversity

-insensitive to sample quality

$H(p, q_\pi) =$	M	$\approx M + \log 2$	$\approx \frac{1}{2}(M - \log \epsilon)$
$H(q_\pi, p) =$	M'	$\approx \frac{1}{2}(M' - \log \epsilon)$	$\approx M'$



reverse Cross-Entropy: $H(q, p) = -\mathbb{E}_{x \sim q_\theta} \log p(x)$

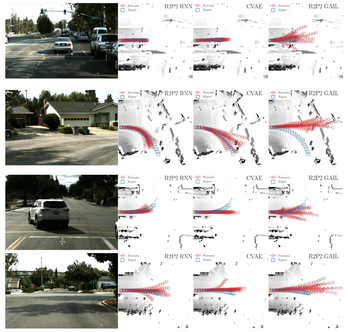
-sensitive to sample quality

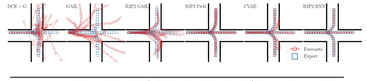
-insensitive to distribution diversity

New public dataset, experiments

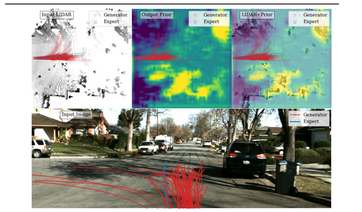
collected new dataset, "CaliForecasting", **3x size of KITTI**

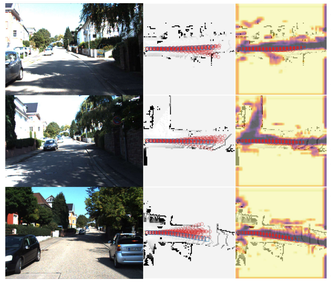
Dataset: http://www.nec-labs.com/~mas/R2P2, Project page: http://www.cs.cmu.edu/~nrhineha/R2P2.html





Approach	DCE-G	GAIL	GAIL†	CVAE [25]	GAIL*	GAIL*†	Linear*	Field*	RNN*
Test $-H(p, q_\theta)$	-0.005	14.10	16.71	-	40.79	36.53	59.20	98.99	**120.7**
Test minMSD	56.90	28.29	21.54	0.001	28.61	1.264	28.54	1.179	**0.0006**





CaliForecasting Approach	Test $-H(p, q_\theta)$	Test minMSD	Test meanMSD	Test $-H(q_\pi, \tilde{p})$
DCE-G	-1.604 ± 0.02	4.953 ± 0.18	11.66 ± 0.27	-129.2 ± 0.43
GAIL-WG [16]	27.43 ± 0.03	9.117 ± 0.27	36.77 ± 2.50	-221.5 ± 2.40
CVAE [25]	$\approx 10.1 \pm 0.9$	1.680 ± 0.12	9.961 ± 0.25	-122.2 ± 0.48
R2P2 GAIL-WG	45.55 ± 0.07	5.529 ± 0.33	25.12 ± 0.80	-152.1 ± 1.00
R2P2 GAIL-WG CNN	43.55 ± 0.08	4.937 ± 0.26	26.59 ± 0.96	-154.3 ± 1.20
R2P2 Linear	64.02 ± 0.11	2.339 ± 0.14	10.51 ± 0.39	-144.5 ± 1.00
R2P2 Linear $\beta = 0.1$	61.57 ± 0.10	2.387 ± 0.13	11.27 ± 0.44	-134.1 ± 0.76
R2P2 Field	54.56 ± 0.11	2.171 ± 0.13	11.59 ± 0.39	-142.5 ± 0.75
R2P2 Field $\beta = 0.1$	53.88 ± 0.11	2.162 ± 0.11	10.87 ± 0.39	-132.8 ± 0.54
R2P2 RNN	$\mathbf{70.20} \pm 0.11$	$\mathbf{1.530} \pm 0.12$	11.25 ± 0.29	-125.0 ± 0.53
R2P2 RNN $\beta = 0.1$	66.89 ± 0.12	1.860 ± 0.14	10.68 ± 0.30	$\mathbf{-119.0} \pm 0.44$
R2P2 RNN $\gamma = 1.0$	65.12 ± 0.12	1.661 ± 0.11	$\mathbf{8.542} \pm 0.22$	-124.8 ± 0.48

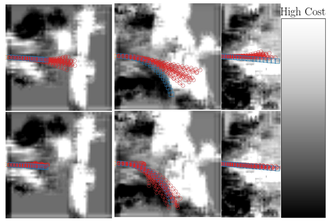
KITTI Approach	Test $-H(p, q_\theta)$	Test minMSD	Test meanMSD	Test $-H(q_\pi, \tilde{p})$
DCE-G	-1.884 ± 0.03	6.217 ± 0.30	15.20 ± 0.62	-137.0 ± 0.72
GAIL-WG [16]	39.53 ± 0.11	5.517 ± 0.34	20.08 ± 2.00	-188.8 ± 1.76
CVAE [25]	$\approx 9.22 \pm 0.9$	1.436 ± 0.15	9.593 ± 0.52	-133.8 ± 1.21
R2P2 GAIL-WG	47.45 ± 0.16	4.062 ± 0.25	13.80 ± 1.10	-168.9 ± 1.50
R2P2 GAIL-WG CNN	42.49 ± 0.12	4.601 ± 0.30	19.87 ± 1.34	-164.2 ± 1.43
R2P2 Linear	62.39 ± 0.14	2.438 ± 0.16	16.16 ± 1.26	-163.4 ± 1.50
R2P2 Linear $\beta = 0.1$	63.82 ± 0.16	2.587 ± 0.15	28.33 ± 1.40	-151.1 ± 1.40
R2P2 Field	64.71 ± 0.18	1.717 ± 0.13	10.34 ± 0.59	-139.2 ± 1.10
R2P2 Field $\beta = 0.1$	62.79 ± 0.29	1.639 ± 0.13	10.92 ± 0.59	$\mathbf{-126.9} \pm 0.77$
R2P2 RNN	$\mathbf{67.70} \pm 0.20$	1.574 ± 0.15	10.46 ± 0.57	-131.6 ± 0.91
R2P2 RNN $\beta = 0.3$	65.80 ± 0.21	$\mathbf{1.282} \pm 0.09$	$\mathbf{9.352} \pm 0.55$	-130.8 ± 0.87

Prior Model

Simple prior approximation with cost map, small number of parameters to further reduce overfitting

$$\min_\theta H(p, \tilde{p}_\theta) = \min_\theta \mathbb{E}_{x \sim p} \sum_{t=1}^{T} \log \frac{e^{-C(x_t, \phi; \theta)}}{\sum_{i \in H} \sum_{j \in W} e^{-C(x_{i,j}, \phi; \theta)}}$$

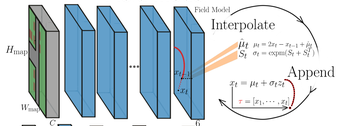
Incorporating prior improves samples



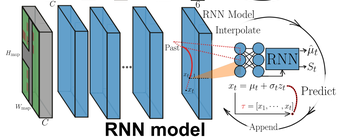
Model learns to perceive obstacles, penalize corresponding positions.

Generator Models

Field model



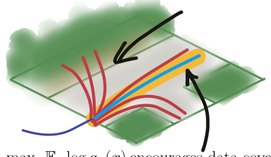
RNN model



Shortcomings of prior work:

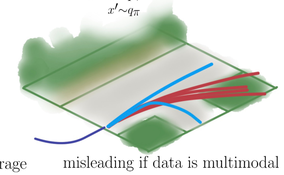
exploitable metrics insufficient for evaluation, no distinction between **sample quality** and **distribution quality**

$\max_\pi \mathbb{E}_{x \sim q_\pi} \log \tilde{p}(x)$ penalizes bad samples



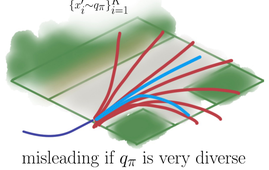
$\max_\pi \mathbb{E}_{x \sim p} \log q_\pi(x)$ encourages data coverage

$\min_\pi \mathbb{E}_{\substack{x \sim p, \\ x' \sim q_\pi}} ||x - x'||^2$



misleading if data is multimodal

$\min_\pi \mathbb{E}_{\substack{x \sim p, \\ \{x'_t \sim q_\pi\}_{t=1}^K}} \min_j ||x - x'_j||^2$



misleading if q_π is very diverse