中国人民大學
RENMIN UNIVERSITY OF CHINA

高瓴人工智能学院
Gaoling School of Artificial Intelligence

Improving Generative Flow Networks with Path Regularization

Anh Do, Duy Dinh, Tan Nguyen, Khuong Nguyen, Stanley Osher, Nhat Ho

University of Texas, Austin
University of California, Los Angeles
FPT Software AI Center

Fanmeng Wang

2023-6-21

Outline

- Introduction
- Background: GFlowNets & Optimal Transport Distance
- Method: Path Regularization via Optimal Transport
- Experiment
- Conclusion

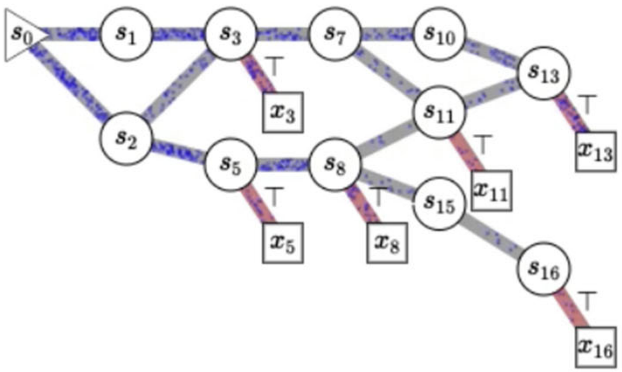
- Introduction
- Background: GFlowNets & Optimal Transport Distance
- Method: Path Regularization via Optimal Transport
- Experiment
- Conclusion

Introduction

➢ **Generative Flow Networks** (GFlowNets) are recently proposed models for learning stochastic policies that generate **compositional objects** by sequences of actions with the probability proportional to a given reward function.

➢ The central problems of GFlowNets are how to improve **exploration** and **generalization.**

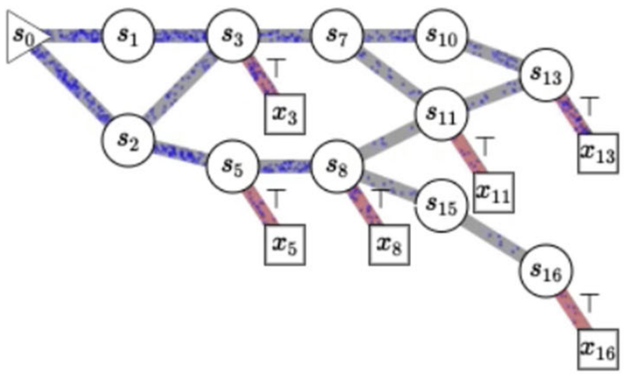
➢ In this paper, we propose to train the GFlowNets with an additional **path regularization** via OT to deal with above problems.



- Introduction
- Background: GFlowNets & Optimal Transport Distance
- Method: Path Regularization via Optimal Transport
- Experiment
- Conclusion

Background

- **GFlowNets: some basic concepts**
  - **Compositional Space *X***, where each object $x \in X$ can be constructed from the source state s_0 and ends in the final state s_f by taking a sequence of discrete actions from the action space A (e.g. a molecule can generated fragment by fragment).



A directed acyclic graph G

node -> state

edge -> action

  - A **complete trajectory τ**, which is a sequence of transitions from s_0 to s_f.

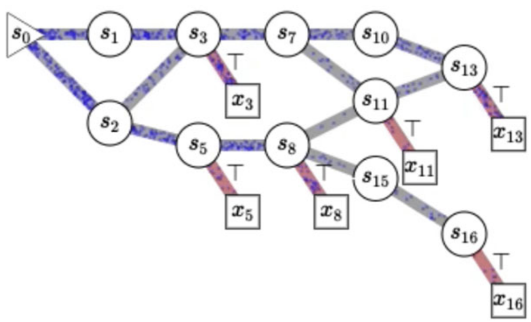
$$\tau = (s_0 \to s_1 \to \ldots \to s_n = x \to s_f)$$

* Bengio E, Jain M, Korablyov M, et al. Flow network based generative models for non-iterative diverse candidate generation[J]. Advances in Neural Information Processing Systems, 2021, 34: 27381-27394.

Background

GFlowNets: Flows

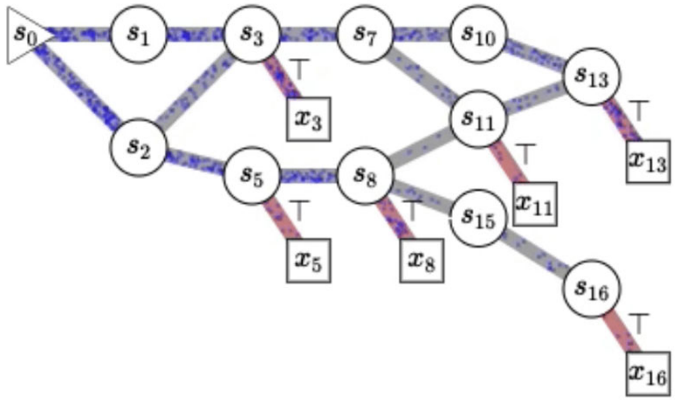
- A **trajectory flow** is a nonnegative function, which represents the probability mass of each complete trajectory τ
- The flow through each state: $F(s) = \sum_{\tau \in \mathcal{T}, s \in \tau} F(\tau)$
- The flow through each edge: $F(s \to s') = \sum_{\tau \in \mathcal{T}, s \to s' \in \tau} F(\tau)$
- The forward transition probabilities(forward policy): $P_F(s'|s) := F(s \to s')/F(s)$
- The backward transition probabilities (backward policy): $P_B(s|s') := F(s \to s')/F(s')$



Background

■ GFlowNets: Learning Objective

- ➢ A GFlowNet can perfectly generate objects proportional to their rewards.
- ➢ In other words, the probability from the source state s_0 to the final state s_f is proportional to the given reward.



Background

■ Optimal Transport Distance

For two discrete probability measures α and β over some space X

- The set of **transportation plans** or **joint probability distributions** can be defined as:

$$\Pi(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \left\{\pi \in \mathbb{R}_+^{k \times l} : \pi \mathbb{1}_l = \boldsymbol{\alpha}, \pi^\top \mathbb{1}_k = \boldsymbol{\beta}\right\}$$

- The Kantorovich **optimal transport** between α and β is defined as follows:

$$\mathrm{OT}_{\mathbf{C}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) := \min_{\pi \in \Pi(\boldsymbol{\alpha}, \boldsymbol{\beta})} \langle \mathbf{C}, \pi \rangle$$

where C is given cost matrix

- Introduction
- Background: GFlowNets & Optimal Transport Distance
- **Method: Path Regularization via Optimal Transport**
- Experiment
- Conclusion

Method

Discrete distance -> OT distance -> Path Regularization

■ Directed distance in the GFlowNet

- We define a new **directed distance** between two arbitrary states in the GFlowNet, which is used as **transportation cost** to compute OT distance.
- Note: the directed distance from a state s to another state s' is designed to be **inversely proportional** to the probability of going from s to s'
- Let τ = (s = s0 → s1 → ... → sn = s′) be the sequence of transitions from s to s′ where $s_t \rightarrow s_{t+1}$ can be a **forward or backward transition** (i.e., the given DAG can be considered an undirected connected graph), the **directed distance** from s to s' is defined as follows:

$$d(s, s') := \min_{\tau=(s\rightarrow\ldots\rightarrow s')} \boxed{-\log(P(\tau \mid s))}$$

i.e., the shortest path from s to s′

The length of the trajectory τ

Method

Discrete distance -> OT distance -> Path Regularization

■ Optimal transport formulation of the path regularization

Give **two neighbor states** s and s′ in trajectory τ , **the forward policy** $P_F(\cdot|\mathrm{s})$ is a discrete probability measure supported by Child(s) = $\{u_1, \ldots, u_k\}$ and $P_F(\cdot|\mathrm{s}')$ is a discrete probability measure supported by Child(s′) = $\{v_1, \ldots, v_l\}$.

➢ The **OT distance** between $P_F(\cdot|\mathrm{s})$ and $P_F(\cdot|\mathrm{s}')$ can be defined as:

$$\mathrm{OT}_{\mathbf{C}}\left(P_F(\cdot|s), P_F(\cdot|s')\right) := \min_{\pi \in \prod(P_F(\cdot|s), P_F(\cdot|s'))} \langle \mathbf{C}, \pi \rangle$$

$$\Pi(P_F(\cdot|s), P_F(\cdot|s')) := \{\pi \in \mathbb{R}_+^{k \times l} : \pi \mathbb{1}_l = P_F(\cdot|s), \pi^\top \mathbb{1}_k = P_F(\cdot|s')\}$$

$$\mathbf{C}_{ij} = c(u_i, v_j) := d(u_i, v_j) = \min_{\tau=(u_i \to \ldots \to v_j)} -\log(P(\tau \mid u_i))$$

➢ The cost matrix C can be calculated in practice by approximating as follows:

$$\begin{tikzcd} s \arrow[r] \arrow[d] & s' \arrow[d] \\ u_i \arrow[r] & v_j \end{tikzcd}$$

$$\mathbf{C}_{ij} = \begin{cases} 0, & \text{if} \quad u_i \equiv v_j \\ \min\left(-\log(P_B(s \mid u_i) P_F(s' \mid s) P_F(v_j \mid s')), -\log(P(v_j \mid u_i))\right), & \text{else if} \quad u_i \to v_j \in \mathcal{A} \\ -\log(P_B(s \mid u_i) P_F(s' \mid s) P_F(v_j \mid s')), & \text{otherwise.} \end{cases}$$

Method

Discrete distance -> OT distance -> Path Regularization

■ Optimal transport formulation of the path regularization

Give **two neighbor states** s and s′ in trajectory τ , **the forward policy** $P_F(\cdot|s)$ is a discrete probability measure supported by Child(s) = $\{u_1, \ldots, u_k\}$ and $P_F(\cdot|s')$ is a discrete probability measure supported by Child(s′) = $\{v_1, \ldots, v_l\}$.

- For any complete trajectory τ , we define the path regularization via OT as follows:

$$\mathcal{L}_{OT}(\tau) := \sum_{t=0}^{n-1} OT_{\mathbf{C}_{t;\theta}}\left(P_F(\cdot|s_t;\theta), P_F(\cdot|s_{t+1};\theta)\right)$$

- If π_θ is the training policy, then **the trajectory loss** is updated:

$$\mathbb{E}_{\tau\sim\pi_\theta}\nabla_\theta(\mathcal{L}_{\mathrm{TB}}(\tau)+\lambda\mathcal{L}_{\mathrm{OT}}(\tau))$$

where λ ∈ R, λ > 0 indicates that we want to **minimize** the path regularization to to improve the generalization and , λ < 0 indicates that we want to **maximize** the path regularization to generate more diverse and novel candidates.

Method

Discrete distance -> OT distance -> Path Regularization

■ Upper bound of optimal transport distance

- Our path regularization's definition requires computing the OT distances for all edges in the trajectory τ , which imposes a heavy burden on the computing
- To overcome this problem, we propose the **upper bound** of the OT distance:

$$
\begin{aligned}
\mathrm{OT}_{\mathbf{C}}\left(P_F(\cdot|s_t), P_F(\cdot|s_{t+1})\right) &:= \min_{\pi \in \prod(P_F(\cdot|s_t), P_F(\cdot|s_{t+1}))} \langle \mathbf{C}, \pi \rangle \\
&\leq \sum_i \sum_j \pi_{ij} \mathbf{C}_{ij} \\
&\leq -\sum_i \sum_j \pi_{ij} \log\left(P_B(s_t|u_i) P_F(s_{t+1}|s_t) P_F(v_j|s_{t+1})\right) \\
&= -\sum_i \sum_j \pi_{ij} \log\left(P_B(s_t|u_i)\right) - \sum_i \sum_j \pi_{ij} \log\left(P_F(s_{t+1}|s_t)\right) - \sum_i \sum_j \pi_{ij} \log\left(P_F(v_j|s_{t+1})\right) \\
&= -\sum_i \log\left(P_B(s_t|u_i)\right) \sum_j \pi_{ij} - \log\left(P_F(s_{t+1}|s_t)\right) \sum_i \sum_j \pi_{ij} - \sum_i \log\left(P_F(v_j|s_{t+1})\right) \sum_j \pi_{ij} \\
&= -\sum_i \log\left(P_B(s_t|u_i)\right) P_F(u_i|s_t) - \log\left(P_F(s_{t+1}|s_t)\right) - \sum_j \log\left(P_F(v_j|s_{t+1})\right) P_F(v_j|s_{t+1}) \\
&= \sum_{u \in \mathrm{Child}(s_t)} P_F(u|s_t) \log(P_B(s_t|u)) - \log(P_F(s_{t+1}|s_t)) + \mathbf{H}(P_F(.|s_{t+1}).
\end{aligned}
$$

Method

Discrete distance -> OT distance -> Path Regularization

■ Upper bound of optimal transport distance

➢ Since the path regularization via OT can be difined as follows:

$$\mathcal{L}_{OT}(\tau) := \sum_{t=0}^{n-1} OT_{\mathbf{C}_{t;\theta}}\left(P_F(\cdot|s_t;\theta), P_F(\cdot|s_{t+1};\theta)\right).$$

➢ The path regularization via OT can be upper bound by:

$$\mathcal{L}_{UB}(\tau) := \sum_{t=0}^{n-1} \left[\sum_{u \in Child(s_t)} P_F(u|s_t) \log(P_B(s_t|u)) - \log(P_F(s_{t+1}|s_t)) + \mathbf{H}(P_F(\cdot|s_{t+1})) \right].$$

$$\sum_{t=0}^{n-1} -\log(P_F(s_{t+1}|s_t)) = -\log\left(\prod_{t=0}^{n-1} P_F(s_{t+1}|s_t)\right) = -\log(P(\tau))$$

- Introduction
- Background: GFlowNets & Optimal Transport Distance
- Method: Path Regularization via Optimal Transport
- Experiment
- Conclusion

Experiment

■ Hyper-grid Environment

This task aims to evaluates the **generalization ability** of the GFlowNet to guess and sample unvisited modes of the interested distribution.

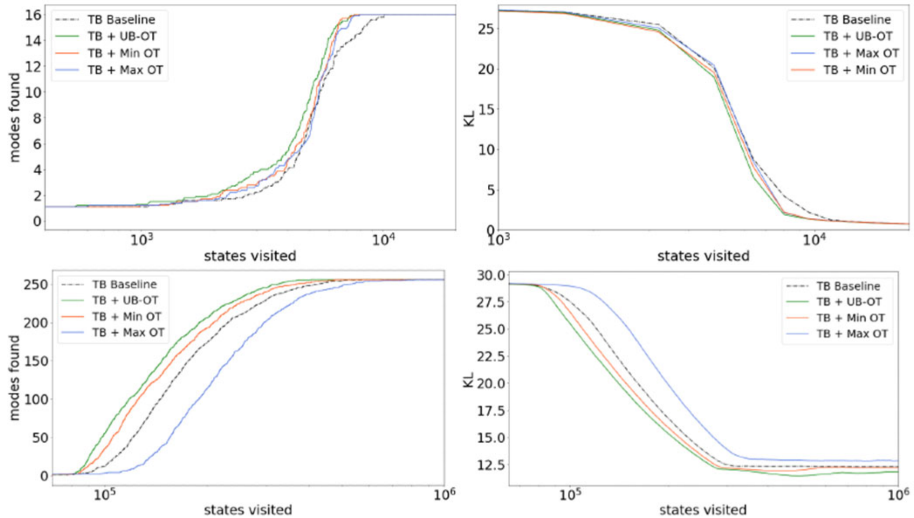


Figure 2: Results on the $4-D$ (upper) and $8-D$ (lower) hyper-grid environment. Left: Number of modes found during training. Right: KL divergence between the true and empirical distribution.

➢ The GFlowNet model trained by **minimizing the path regularization via OT** and **the upper bound** both performer well in improving generalization ability.

高瓴人工智能学院
Gaoling School of Artificial Intelligence

Experiment

■ Synthetic Discrete Probabilistic Modeling tasks

➢ Training the GFlowNet with either **minimizing** the path regularization via OT (Min OT) or via the **upper bound** (UB OT) gains the better **NLL** and **MMD** scores than the baseline and Max OT.

Metrix	Method	2spirals	8gaussians	circles	moons	pinwheel	swissroll	checkerboard
NLL ↓	PCD	20.094	19.991	20.565	19.763	19.593	20.172	21.214
	ALOE	20.295	20.350	20.565	19.287	19.821	20.160	54.653
	ALOE +	20.062	19.984	20.570	19.743	19.576	20.170	21.142
	EB-GFN (paper)	20.050	19.982	20.546	19.732	19.554	20.146	20.696
	EB-GFN	20.0679	19.9862	**20.5598**	19.7324	19.5735	20.1599	20.6839
	EB-GFN + Max OT	20.0673	19.9857	20.5599	19.7319	19.5714	20.1597	20.6837
	EB-GFN + UB OT	20.0651	**19.9854**	20.5600	**19.7305**	19.5707	20.1596	20.6836
	EB-GFN + Min OT	**20.0640**	19.9855	**20.5598**	19.7308	**19.5699**	**20.1595**	**20.6831**
MMD ↓	PCD	2.160	0.954	0.188	0.962	0.505	1.382	2.831
	ALOE	21.926	107.320	0.497	26.894	39.091	0.471	61.562
	ALOE +	0.149	0.078	0.636	0.516	1.746	0.718	12.138
	EB-GFN (paper)	0.583	0.531	0.305	0.121	0.492	0.274	1.206
	EB-GFN	0.3012	0.0408	−0.1724	−0.1744	0.2056	0.1555	−0.0986
	EB-GFN + Max OT	0.3258	0.0197	−0.1919	−0.0456	0.1377	0.0763	−0.0903
	EB-GFN + UB OT	0.2902	**0.0102**	**−0.2819**	−0.1253	0.1561	**0.0257**	−0.0923
	EB-GFN + Min OT	**0.1816**	0.0343	−0.2775	**−0.1966**	**0.1220**	0.1334	**−0.1071**

Table 4: Results on the Synthetic EB-GFN tasks. The negative log-likelihood (NLL) and MMD are displayed in units of 1×10^{-4}. ALOE+ uses a 30 larger parametrization than ALOE and EB-GFN. We only take into account the the reproduce results of EB-GFN when comparing with our methods (Min OT, Max OT, and UB OT).

* Zhang D, Malkin N, Liu Z, et al. Generative flow networks for discrete probabilistic modeling[C]//International Conference on Machine Learning. PMLR, 2022: 26412-26428.

Experiment

■ Biological Sequences Design

- The experiments are conducted in the multi-round active learning setting, with the goal of **generating a diverse set of useful candidates** after evaluation rounds.

	Performance	Diversity	Novelty
DynaPPO	0.938 ± 0.009	12.12 ± 1.71	9.31 ± 0.69
COMs	0.761 ± 0.009	19.38 ± 0.14	26.47 ± 1.30
GFlowNet-AL (paper)	0.932 ± 0.002	22.34 ± 1.24	28.44 ± 1.32
GFlowNet-AL	0.874 ± 0.022	$\mathbf{31.98 \pm 2.27}$	23.91 ± 1.87
GFlowNet+Min OT-AL	0.847 ± 0.033	20.32 ± 7.38	23.63 ± 1.66
GFlowNet+UB OT-AL	0.828 ± 0.022	29.89 ± 2.80	24.16 ± 1.75
GFlowNet+Max OT-AL	0.917 ± 0.003	31.56 ± 2.43	$\mathbf{28.86 \pm 0.96}$

Table 1: Results on the AMP task with $K = 100$.

	Performance	Diversity	Novelty
DynaPPO	0.58 ± 0.02	5.18 ± 0.04	0.83 ± 0.03
COMs	0.74 ± 0.04	4.36 ± 0.24	1.16 ± 0.11
BO-qEI	0.44 ± 0.05	4.78 ± 0.17	0.62 ± 0.23
CbAS	0.45 ± 0.14	5.35 ± 0.16	0.46 ± 0.04
MINs	0.40 ± 0.14	$\mathbf{5.57 \pm 0.15}$	0.36 ± 0.00
CMA-ES	0.47 ± 0.12	4.89 ± 0.01	0.64 ± 0.21
AmortizedBO	0.62 ± 0.01	4.97 ± 0.06	1.00 ± 0.57
GFlowNet-AL (paper)	0.84 ± 0.05	4.53 ± 0.46	2.12 ± 0.04
GFlowNet-AL	0.83 ± 0.01	4.66 ± 0.08	1.14 ± 0.03
GFlowNet+Min OT-AL	0.82 ± 0.01	4.72 ± 0.10	1.13 ± 0.04
GFlowNet+UB OT-AL	0.83 ± 0.01	4.68 ± 0.10	1.14 ± 0.05
GFlowNet+Max OT-AL	$\mathbf{0.85 \pm 0.02}$	4.52 ± 0.18	$\mathbf{1.21 \pm 0.10}$

Table 2: Results on the TF Bind 8 task with $K = 128$.

* Jain M, Bengio E, Hernandez-Garcia A, et al. Biological sequence design with gflownets[C]//International Conference on Machine Learning. PMLR, 2022: 9786-9801.

Experiment

■ Biological Sequences Design

- The experiments are conducted in the multi-round active learning setting, with the goal of **generating a diverse set of useful candidates** after evaluation rounds.

	Performance	Diversity	Novelty
DynaPPO	0.794 ± 0.002	206.19 ± 0.19	203.20 ± 0.47
COMs	$\mathbf{0.831 \pm 0.003}$	204.14 ± 0.14	201.64 ± 0.42
BO-qEI	0.045 ± 0.003	139.89 ± 0.18	203.60 ± 0.06
CbAS	0.817 ± 0.012	5.42 ± 0.18	1.81 ± 0.16
MINs	0.761 ± 0.007	5.39 ± 0.00	2.42 ± 0.00
CMA-ES	0.063 ± 0.003	201.43 ± 0.12	203.82 ± 0.09
AmortizedBO	0.051 ± 0.001	205.32 ± 0.12	202.34 ± 0.25
GFlowNet-AL (paper)	0.853 ± 0.004	211.51 ± 0.73	210.56 ± 0.82
GFlowNet-AL	0.8232 ± 0.0001	218.54 ± 7.88	222.05 ± 5.49
GFlowNet+Min OT-AL	0.8231 ± 0.0001	182.03 ± 0.25	220.56 ± 2.02
GFlowNet+UB OT-AL	0.8232 ± 0.0001	221.64 ± 0.11	218.02 ± 0.79
GFlowNet+Max OT-AL	0.8233 ± 0.0001	$\mathbf{225.00 \pm 3.76}$	$\mathbf{242.11 \pm 1.44}$

Table 3: Results on the GFP task with $K = 128$.

- The GFlowNet-AL model trained by **maximizing the path regularization via OT** performs well in generating more diverse and novel candidates.

- Introduction
- Background: GFlowNets & Optimal Transport Distance
- Method: Path Regularization via Optimal Transport
- Experiment
- Conclusion

Conclusion

- We propose to train the GFlowNet with **an additional path regularization via Optimal Transport** that places prior constraints on the underlying structure of the GFlowNet.
- We derive **an efficient implementation** of the regularization by finding its **closed-form solutions** in specific cases and a meaningful **upper bound** that can be used as an approximation when we want to minimize the regularization term.
- Experiments have shown that **minimizing the path regularization** via OT improves the GFlowNet's **generalization** while **maximizing the path regularization** via OT enhances the **exploration ability** of the GFlowNet.

中国人民大學
RENMIN UNIVERSITY OF CHINA

高瓴人工智能学院
Gaoling School of Artificial Intelligence

Thank You for listening!

Fanmeng Wang

2023-6-21