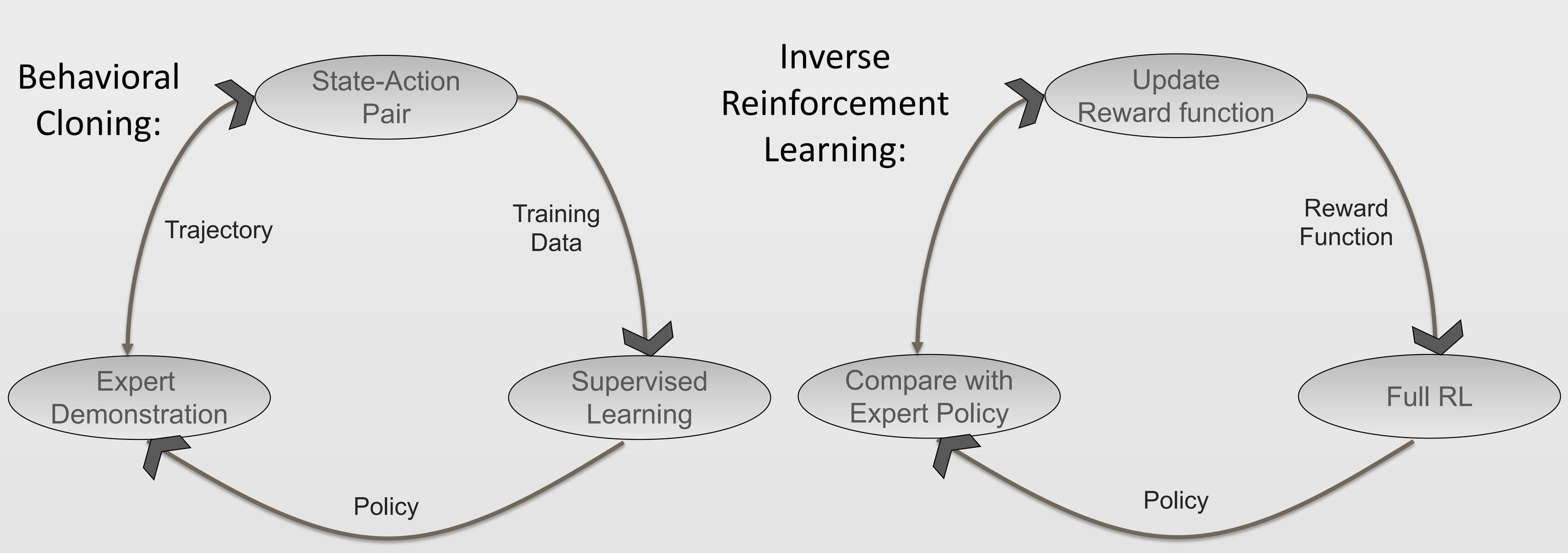
Imitation Learning for Medical Dosage Prescription

Motivation

Imitation learning techniques aim to mimic human behavior in a given task. An agent (a learning machine) is trained to perform a task from demonstrations by learning a mapping between observations and actions. It is widely applied in areas like self-driving car and. We are considering applying this idea to prescription such that agent will learn optimal dosage policy.

There are various imitation learning methods, one of the most well-known and simplest to implement one is **behavioral cloning**, a method of learning from expert trajectories by dividing them into state-action pairs. However, with the assumption of i.i.d, errors made in different states add up, and a mistake made by the agent can easily drift away from demonstrated states. Such accumulated error may cause severe problem in medical dosing cases.

There are also methods based on **reinforcement learning**, like inverse reinforcement learning. It's solving the error problem of BC, but it's reward function of environment need to be learned based on expert demonstrations, which is very complex.



Therefore, we looked at **Soft Q imitation learning**. It is also adopting part of the idea of reinforcement learning like IRL, while it is having **constant reward function** instead. With this method, the problem of accumulated error and complex training process is avoided, and it is especially appropriate for application to prescriptions.

Objective

The main focus of this project is to apply imitation learning with idea of reinforcement learning with constant reward to medication prescribed dosage data:

- learn optimal prescription policy.
- improve the sample-efficiency of the algorithm

Methodology

The main algorithm we will adopt is soft Q imitation learning, which is an algorithm instantiated with soft Q-learning. We can initialize the agent's experience replay buffer with expert demonstrations, setting the rewards to a **constant**, r = +1 in the demonstration experiences, and setting rewards to a constant r = 0 in all of the new experiences the agent collects while interacting with the environment.

Algorithm 1 Soft Q Imitation Learning (SQIL)

1: Require $\lambda_{\text{samp}} \in \mathbb{R}_{\geq 0}$
2: Initialize $\mathcal{D}_{\text{samp}} \leftarrow \emptyset$
3: **while** $Q_{\boldsymbol{\theta}}$ not converged **do**
4: $\quad \boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \eta \nabla_{\boldsymbol{\theta}} (\delta^2(\mathcal{D}_{\text{demo}}, 1) + \lambda_{\text{samp}} \delta^2(\mathcal{D}_{\text{samp}}, 0))$ {See Equation 1}
5: $\quad$ Sample transition (s, a, s') with imitation policy $\pi(a|s) \propto \exp(Q_{\boldsymbol{\theta}}(s, a))$
6: $\quad \mathcal{D}_{\text{samp}} \leftarrow \mathcal{D}_{\text{samp}} \cup \{(s, a, s')\}$
7: **end while**

For the sample efficiency, there several potential methods.

- Off-policy, enabling more efficient sample re-use.
- self-supervised auxiliary losses

Dataset and Application

The dataset that we will apply our algorithm to is **prescribed medicines**. It is extracted from the 2017 Medical Expenditure Panel Survey (MEPS).

Variable	Description
RXBEGMM	Month person first used medicine
RXBEGYRX	Year person first used medicine
RXNAME	Medicine name (Imputed)
RXDRGNAM	Multum medicine name (Imputed)
RXNDC	NDC (Imputed)
RXQUANTY	Quantity of Rx/prescribed medicine (Imputed)
RXFORM	Dosage form (Imputed)
RXFRMUNT	Quantity unit of medication (Imputed)
RXSTRENG	Strength of medication (Imputed)
RXSTRUNT	Unit of medication (Imputed)
RXDAYSUP	Days supplied of prescribed med(Imputed)

The data includes medicine name, the month and year the person first use the medicine, quantity of prescribed medicine, dosage form, etc. These information are important to multiple public health reporting and research use cases to track and analyze information about treatment for various health conditions. The dosage of medicines prescribed is what we mainly want to learn with SQIL

References

Reddy, Siddharth, Anca D. Dragan, and Sergey Levine. "SQIL: Imitation Learning Via Reinforcement Learning with Sparse Rewards." (2019).

Hussein, Ahmed, Mohamed Gaber, Eyad Elyan, and Chrisina Jayne. "Imitation Learning: A Survey of Learning Methods." ACM Computing Surveys 50, no. 2 (2017;2018;): 1-35.

Zoltán Lőrincz. "A Brief Overview of Imitation Learning." (2019).