A Survey of Robot Learning from Demonstration

Authors: Brenna D.Argall, Sonia Chernova, Manuela Veloso and Brett Browning.
Year: 2009
Journal: Robotics and Autonomous Systems

Report:

1. **Abstract:**

"We present a comprehensive survey of robot Learning from Demonstration (LfD), a technique that develops policies from example state to action mappings. We introduce the LfD design choices in terms of demonstrator, problem space, policy derivation and performance, and contribute the foundations for a structure in which to categorize LfD research. Specifically, we analyze and categorize the multiple ways in which examples are gathered, ranging from teleoperation to imitation, as well as the various techniques for policy derivation, including matching functions, dynamics models and plans. To conclude we discuss LfD limitations and related promising areas for future research."

Introduction

The paper intends to deliver the methods and nomenclature of the Learning from Demonstration(LfD) from scratch. The procedures from initializing input to the extracting output and their combinations are discussed. The key design parameters, approaches to LfD and core techniques for policy derivation and improving the optimization are also been concentrated.The problem is divided into two major problems as:

- Gathering the information
- Deriving the policy

Content

The classical approaches for Robotic learning are based on the complexity of the world model. For that reason, even the experts have to linearize the problem definition, which decreases the approximation and precision of the solution.

- For this reason, the reinforcement learning with feedback rewards are adapted to increase the efficiency of the demonstration.
- The approach starts with selecting the demonstrator type and technique.
- Demonstrator can be chosen based on two keypoints:

```
1.          Who          controls          the          demonstration
 2.          Who          executes          the          demonstration
```

- Next we have to choose whether the choice of action is continuous or discrete.
- Generally the core approaches for policy derivations are classified into three main types, which we have to chose one of three:

1. **Mapping Function**

- In mapping function, generally the policy derivation is not done.
- It only allows an generalization of solution based on demonstration.
- Hence we can note that, it only approximates the solution which is not precise.

2. **System Model**

- It collects the data from the world, analyzes and select a policy based on the world dynamics.
- It deals with the transition function , which assists in deriving the policy.
- A reward function is also used by the process to increase the efficiency of the output. It is also further derived into two based on input of the reward function:

1. Engineered Reward Functions
2. Learned Reward Functions

3. **Plans:**

- In this process of deriving policy, specific actions are defined for the

specific conditions.

- So, it reads the pre-condition and act accordingly to result in post

condition.

- The correspondence is a term which deals with the mapping of the data from the

input to the output through mappings. For that reason, two types of mappings are introduced.

1. Record mapping
2. Embodiment mapping

- Each one of the following as shown in the figure can be described as one of the

example with the identical mapping and some function mapping.

- **Demonstration:** In this process, the embodiment mapping is identical, (no function for embodiment) and it is performed directly on the robot learner.
- **Imitation:** As the name defines, it depicts that the learner imitates the teacher, so it requires the function to project the imitations. So embodiment mapping is used.

Demonstration is further divided into two types as:

1.**Teleoperation:** It is a technique in which the both embodiment and record mapping are absent.It is the process of operating robot with the help of joystick and it remembers it’s own actions and perform the same when needed.

2.**Shadowing:** This technique is an example for the absent of embodiment mapping and use of record mapping. This can be explained as imitating the

teacher by sensors of robots. The recorded version of teaching is sent in a way that robot sensors can illustrate the teaching.

● Imitation is also further divided into two types based on the presence and absence of the record mapping.

1. **Sensors on teacher:** This is a technique in which the sensors are on the teachers, and the the actions are recorded using the sensors. Here the record mapping is absent.

2. **External Observation:** Here the record mapping is done. In this the external sensors execute the actions of learner by the input of the teacher.

● The classification given are just the examples, in which many of such cases may appear.

Discussion and Thoughts

The paper also concentrated on stabilizing the terminology. As before many papers worked with different terminology for the same procedure, they tried to use a balanced terminology.

- A good example of pick and place of robot is taken as instance. As it is simple and more ideal, they covered many topics with the same example.
- Various definitions were also noted along the paper, which are very important.
- The difficulties of following different procedures and the ways to overcome are also described.
- The different papers related to corresponding topics well referenced which are very required. On the overall, it appeases to be a survey paper

Conclusion

- A good overview of robot control through various procedures and their classification is described in the article appealingly.
- There have been also the comparisons done in between the algorithms and their corresponding pros and cons have been explained briefly
- The paper described the methods of LfD very effectively. The steps of the algorithm of the LfD at different stages were also well defined and explained.

2. **Sources:**

[1] https://www.sciencedirect.com/science/article/pii/S0921889008001772