NAME- VEDANT PULAHRU

COURSE – COMP3190

STUDENT NUMBER - 7876784

PROFESSOR - Dr. John Anderson

Imitation Learning: Learning and Replicating actions/gestures by Imitation

INTRODUCTION TO THE TOPIC

Robotics is the combined study of science and technology and the technology that can perform this, are called Robots. These robots are somewhat like newborn babies, i.e., they are learning by exploring their environment or in better terms, human guidance. Imitation learning is a technique where the knowledge is transferred by mimicking a human or a certain behavior for any given task. In the past few years, the demand for imitating human behavior has increased which resulted in the rise of many aspects that need the concept of artificial intelligence. The Robot Programming by Demonstration (PbD) covers the method through which robots learn new skills with human guidance. PbD covers a broad range of applications. For example, assistive robots such as industrial robots, where the goal is to reduce the time and cost to program the robot. Pbd would create several versions of this product and then the work can be done without the help of an expert in robotics. Another example of AI is self-driving cars, where the AI of the car reads the road marking and using google maps, can take the passenger from one destination to another. Self-driving cars are pre-coded models where they only perform a certain function rather than imitating someone’s behavior and learn by it.

The field of imitation learning draws its importance from its relevance to a variety of application such as human computer interaction [Hussein, Ahmed, 2017]. This is used to teach the robots about varying skeletons and degree of freedom (DOF) to perform different tasks. For example, for a humanoid robot, we need a high degree of freedom (DOF) so that it can learn discrete actions such as walking (cyclic tasks), standing up (simple task).

We are also not that different from these robots. Humans are learning from the time they were babies by observing people around them. Infants are constantly experimenting to determine their capability [Allen, Jeff, and John Anderson, 2010]. Expressions and body movement along with body behavior are some of the things that an infant tends to learn/copy from the people around them. As they grow up, learning expands. A human naturally deals with heterogeneous demonstrators: if a child's first exposure to the game of frisbee is through observing a dog catching a frisbee in its mouth, then the child will try catch it with their hands instead and not using their mouth. This way they learn the task using the skills that are natural and available to them regardless of the demonstration they learned from [Allen, Jeff, and John Anderson, 2010].

The imitation process starts by sensing the actions, via different sensing methods. After sensing the action, the data is processed to emulate the action which was described by the performer. As the usage of this technology is increasing, the future is driven towards relying on the ability of artificial intelligence due to which, more problems arise. A fundamental problem when it comes to Imitation learning is to create an appropriate mapping between actions by the demonstrator and the machine. When it comes to complex applications, the number of possible scenarios is too large to cover. While this is formulated as an optimization problem, one of the major drawbacks is that they rely on a large amount of prior knowledge, due to which more inefficiency arises.

A research report called "Imitation learning: A Survey of Learning Methods" by Ahmed Hussein, et al stated some of the major problems faced in imitation learning due to its interdisciplinary nature which results in increase in inefficiency in the output even if the model is well prepared. It includes:

- Correspondence problem: matching of learner's capability, degree of freedom and skeleton of the demonstrator. Any difference in size or structure between learner and teacher must be compensated during training.
- Observability problem: this means if the demonstration is not provided by the demonstrator, the learner may not be aware of the capabilities and possible action of the human.
- The policy must be able to adapt to variations in the task and the environment. The complex nature of imitation learning applications dictates that agents must be able to perform the task correctly even under different scenarios that are different from demonstrations
- The Noisy and unreliable sensing, and correspondence of the machine and the human (the teacher), Including few more which resulted in inefficiency.

Learning a direct mapping between the state and action is not often enough to reach or achieve the required outcome/behavior. This can happen due to numerous amounts of issues such as errors in understanding the demonstration, variance in the skeletons of the demonstrator and robot. Moreover, humans' ways of performing task might vary from the robot because of different environment.

HIDDEN MARKOV MODEL

Hidden Markov Models (HMM) can be used to recognize the primitive sequence in the visual data gathered from demonstration by others [Allen, Jeff, 2009]. The steps include:

Visual data collection -> Representing it for the purpose of training the model and -> Implementing as well as training the HMM.

The model which we will discuss later, will be a core representation of Hidden Markov Model (ICub). We will compare approach with four other methods [Calinon, Sylvain]:

1) TGMR: Time-dependent gaussian mixture regression

   In this method, time is used as explicit input variable. The distribution of temporal and spatial variables is encoded in Gaussian Mixture Model (GMM). The controller used by the robot to reproduce the skill is the mass-spring damper system

2) LWR: Locally Weighted Regression

   This is a memory based probabilistic approach. It is used to estimate the desired position and velocity. The controller used by the model is a mass-spring damper system.

3) LWPR: Locally Weighted Projection Regression

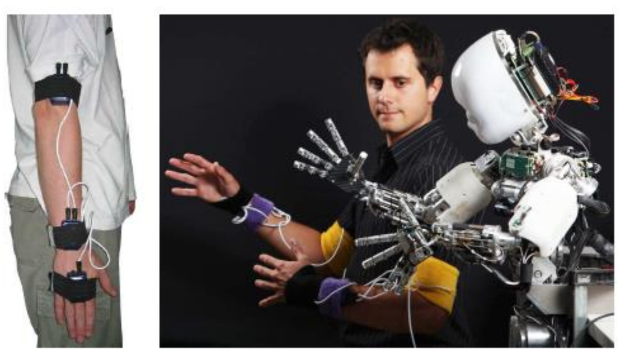
   This is an incremental regression algorithm that performs piecewise linear function approximation. This algorithm does not need to store the training data and has been proved to be efficient in robot learning including high dimensional data. This method reduces dimensionality of the input data by finding local projection through Partial Lease Squares regression. The controller used by the robot is the mass-spring damper system.

4) DMP: Dynamic Movement Primitive

   This method allows to reach the goal by modulating a set of mass-spring-damper system.

EXPERMIENT 1: ICUB (HUMANOID ROBOT)

Multiple models have been created to understand this concept. For example- ICub is a humanoid robot from the European project RobotCub. To understand the movement, sensors have been used to reproduce what the performer performed. Each sensor provides 3D absolute orientation of each segment by integrating the 3D rate of turn, acceleration, and earth-magnetic field at a rate of 50 Hz and with a precision of 1.5 degree. The upper torso is defined as a kinetic chain where the shoulder joint connects the gridle and the upper arm, the elbow connects the upper arm and forearm, the wrist connects the forearm and the hand [Calinon, Sylvain].



(Left: motion sensors), (right: imitator trying to imitate the demonstrator)

CONCLUSION OF EXPERIMENT 1:

In this experiment, we found that the high dimensional periodic movement with crossing was correctly handled by DMP and HMM. DMP showed the best score in terms of smoothness and accuracy. In contrast, LWR and LWPR couldn't handle the crossing point correctly during movement. From an algorithmic point of view, passing through the same point several times along with motion cannot be handled by LWR and LWPR [Calinon, Sylvain]. When reaching the crossing points, both methods provided inadequate motion behavior. Similarly, with TGMR method, we couldn't efficiently encode the periodic motion due to explicit encoding of time in the model.

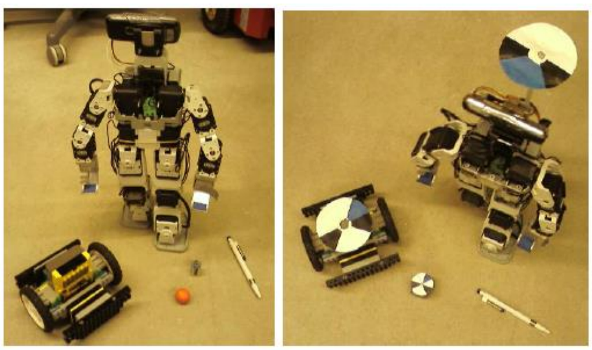
EXPERMENT 2: TWO-WHEELED DRIVE ROBOT

Second example that we are going to use is a two-wheeled differential-drive robot which is from [Allen, Jeff, 2009]. We have a total of 3 robots, one is a robot imitator, which is two wheeled differential-drive robot. One of them is used as a demonstrator which is physically identical to the imitator to compare how well the imitator learns from heterogeneous demonstrator. the first is a humanoid robot based on Bioloid Kit because it provides a different physiology from imitator in terms of how motion made by robot appears visually. The third robot is two wheeled Citizen Eco-Be robot which is 1/10 the size of the imitator. This bot was chosen because the size difference and difficulty in moving a ball due to light weight. The robot learns by observing one demonstrator at a time where the demonstration is to shoot a ball into an empty goal. Furthermore, there all knowledge was gained by observing and no communication between the imitator and demonstration.

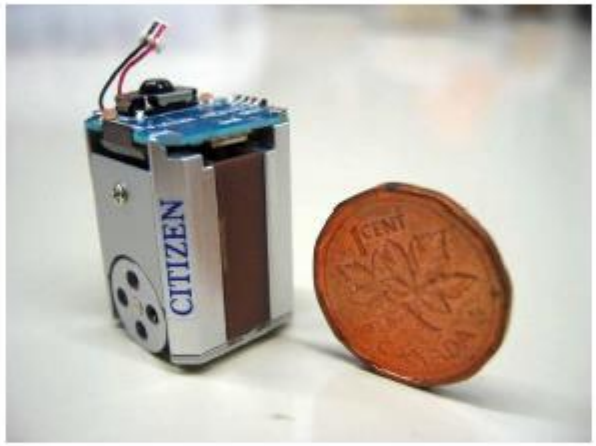
There are 5 atomic motor commands available to the wheeled imitator which are

1. Forward
2. Backward
3. Left
4. Right
5. Stop

As we know, in imitation learning, the imitator collects visual data from the action to create an understanding of what it should do. Behaviors are learned by combining primitive or existing behavior to produce more complex actions based on observation. Behavior is built and stored using a type of forward model (which is one of the robots we were using) which represents frequencies of primitive and behavior occurring in sequence. Also, they are used to explain and predict the behavior of demonstration in terms of imitator's repertoire.



(Two views of heterogenous robot)



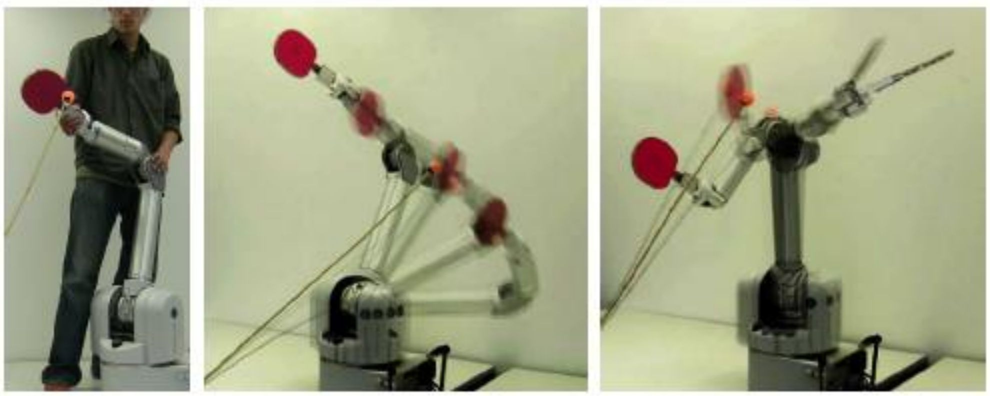
(Close view of Citizen Eco-Be microrobot)

CONCLUSION OF EXPERIMENT 2:

To determine if the imitator was exposed to the demonstrations and learned, we ordered demonstration in two ways. In first, we have Mindstorms robot demonstrator, then the Citizen demonstrator and finally we have Bioloid demonstrator (in short RCB). The second ordering is the reverse of this i.e., Bioloid, Citizen and finally Mindstorms (in short BCR). This ordering did not affect the total number of behaviors created or deleted from the forward models. The statistical method used to calculate the preconditions were not robots and had a small sample size to work with which resulted in the robot driving closer to the ball. But this is fixable with more precise preconditions and possible initial training of the imitator.

EXPERIMENT 3: WAM ROBOTIC ARM

In the third example, we are going to talk about WAM ROBOTIC ARM (image given below) [Calinon, Sylvain]. In this experiment, they aim to demonstrate tat framework can be used in an unsupervised learning manner. In this experiment the robot arm will learn and reproduce hitting a ball with a table tennis racket. They also increased the difficulty of this experiment by assuming that robot can hit the ball with any desired velocity from retrieved from the demonstrator.



(Left: setting up the experiment), (middle: producing drive stroke), (right: producing topspin)

In table tennis, topspin happens when the top of the ball is going in the same direction as the ball then, the ball drops faster than the gravity and hit the table to get a point. Drive is a stroke with no spin and the motion is different in both topspin and drive stroke. The robot was trained by demonstrating several of these by an intermediate-level player. The player demonstrated a total of 4 topspin and 4 drive strokes in a random order where what type of move is this, was not provided by implementer. The number of states in the HMM is selected through Bayesian Information Criterion (BIC).

CONCLUSION OF EXPERIMENT 3:

In the result of this experiment, they saw that HMM approach reproduces an appropriate amount of motion in the two situations. The model was able to learn these two different dynamics which tells us that the model can encode several other motion alternatives along with the two stated above. The robot arm was correctly able to attain the ball coming at a certain velocity, which was like the one demonstrated.

CONCLUSION

With the advancement in technology, we certainly have discovered different ways of performing different types of learning processes. But still, there are many loopholes when it comes to the world of robotics. We presented an evaluation of how imitation in ICub, heterogenous robots' example as well as Wam robotic arm worked. In all these scenarios, we did see good results but in some, we couldn't achieve a "close to perfect model". To achieve a "one-shot imitation" we would have to improve the efficiency and data. The imitation learning concept is still expanding and over time, we will surely have better models. But with those models, we will have even bigger and complex data and tasks to complete its concept.

BIBLIOGRAPHY

- Calinon, Sylvain, et al. “Learning and Reproduction of Gestures by Imitation.” IEEE
- Robotics & Automation Magazine, vol. 17, no. 2, IEEE, 2010, pp. 44–54, doi:10.1109/MRA.2010.936947.
- Hussein, Ahmed, et al. “Imitation Learning: A Survey of Learning Methods.” ACM Computing Surveys, vol. 50, no. 2, ACM, 2017, pp. 1–35, doi:10.1145/3054912.
- Allen, Jeff, and John Anderson. “Heterogeneous Imitation Learning from Demonstrators of Varying Physiology and Skill.” *2010 Ninth International Conference on Machine Learning and Applications*, IEEE, 2010, pp. 105–12, https://doi.org/10.1109/ICMLA.2010.23
- Allen, Jeff. *Imitation Learning from Multiple Demonstrators Using Global Vision*. 2009.