Robot Learning from Observations

Abstract: In this article we present a view that robot learning should make use of observational data in addition to learning through interaction with the world. We hypothesize that the acquisition of many, if not the majority, of features of intelligent robots can happen or be drastically accelerated through learning by observation. This view is in part motivated by extensive studies on observational learning in cognitive science and developmental psychology. There are, however, many aspects of observational learning. To help facilitate future progress, we construct a taxonomy of observational learning. At the highest level of the hierarchy, we categorize it into learning representations, models of the world, behaviors, and goals. Thanks to the internet and advances in augmented reality hardware, we may soon be able to observe large fractions of everyday human activities and interactions with the world, which puts us in a good position to develop robots that learn to act in the context of rich observational experience.

Keywords: Robot Learning, Observational Learning

1 Introduction

The dream of robotics has always been that of general purpose machines that can perform many diverse tasks in unstructured human environments. One way to arrive at such machines is to embrace learning from ground up. The most popular paradigm for this is currently reinforcement learning (RL), which involves trial and error learning through interaction with the environment [1]. We have recently witnessed impressive results in RL, achieved by large scale training in simulation across an ensemble of randomized environments [2]. In current instantiations, such techniques suffer from three main limitations: large sample complexity, considerable human effort in reward function design, and limited generalization ability.

Consider a child learning to play tennis. Even though the child has never played the game of tennis, it comes equipped with rich representations and models of the world. Thus, the child is able to identify the racket and the balls, knows that a racket is held at the handle, and that if the ball collides with the net it will stop. Furthermore, the child has also likely observed others play tennis. Consequently, it has a good sense of behaviors required to hit the ball and is able to roughly reproduce them, and knows that the goal of the game involves getting the ball onto the opponent’s side of the court.

While pedagogical, the tennis example hints at ways to improve our current robot learning systems. In particular, the key message is that the child has learnt a lot about the world and the game of tennis through observation before even starting to play. Thanks to this, it does not need to discover everything through interaction and is able to learn to play much quicker. Likewise, our robots should make use of observational data in addition to learning through interaction with the world.

Thanks to the internet, we now have access to a large amounts of video data from which we may hope to learn in various ways. We can observe physical phenomena such as apples falling, balls bouncing, and water flowing. We can observe humans performing cartwheels, juggling, baking cakes, playing soccer and assembling furniture. Thus "observational learning" embraces several different capabilities - acquiring commonsense physics, rules of social behavior, and procedural knowledge such as how to change a tire.
A very important class of learning, and one at which normal human children are particularly adept at, is “social learning”, well studied in cognitive science and developmental psychology. It begins early, e.g. facial imitation in children [3, 4, 5]. When compared with primates, human babies are comparable at motor control but excel at social learning [6]. More generally, social learning can be seen as a core component of cultural transmission of knowledge across generations [7, 8]. In a similar spirit, social learning theory [9] argues that learning is not purely behavioral but takes place in a social context. Thus, learning can occur by observing behaviors of others acting in the world.

We focus particularly on observational data that comes in the form of videos, optionally with sound, text, or metadata. Crucially, this data should be recorded in the wild and be as representative of the real world as much as possible. The nature of such data, may vary along a number of dimensions. We highlight a difference between third-person data, recorded by an observer (top left), and first-person data, recorded by the camera wearer (bottom left). Majority of the observational data currently available on the internet is third-person. This is a side-effect of the types of camera devices we use, most commonly phones, and the “interestingness” bias that people have when uploading data to the web. With the advances in augmented reality, we may soon get access to endless streams of first-person data of daily activities of people and their interactions with the world. These may include boring everyday things that are perhaps not of interest to the internet but may prove useful for robotics. If we make an analogy with camera images and how their wide availability has impacted computer vision, it might be that this sort of data of human activities proves particularly crucial.

The contribution of this paper is to construct a taxonomy of observational learning. At the highest level, we categorize it into learning representations, models, behaviors, and goals.

## 2 Observational Learning

In the previous section, we presented a case for robot learning from both observational and interaction data. But what might we be able to learn from observational data?

Here, we propose a taxonomy of observational learning and give concrete examples. At the highest level of the hierarchy, we can learn representations, models, behaviors, and goals (Figure 2).

### 2.1 Learning Representations

When a child is faced with learning a new task, like in the tennis example, it does not start by seeing an unstructured collection of pixel values. But rather, it is already equipped with rich representations of the world. These allow for understanding objects, parts, affordances, and many other aspects. Most importantly, these have largely been acquired through observational learning.

Likewise, our robots could leverage observational data to build rich representations. Such representations would allow them to learn to act in the context of prior experience of the world. These representations further facilitate learning models, behaviors, and goals on top, discussed below.

While starting from pixels is one extreme, operating directly on top of state representations such as objects with known 6 DoF pose, is the other extreme. Such representations are difficult to acquire and can lead to compounding errors when used directly. There is a range of possibilities in between.
Computer vision community has now developed reliable systems for detecting objects [10, 11] that could serve as useful representations, e.g. for building dynamics models [12]. More specific than objects are object part keypoints [13] and even more specialized than that are representations of humans [14]. We might also want to capture various scene properties, such as depth or surface normals, which can help learning policies in navigation [15] or manipulation [16].

While many of the examples listed above use supervision to arrive at such representations these could, and in the longer term should, be learnt in an unsupervised way [17, 18, 19, 20, 21]. Indeed, there is already evidence of such representations being useful for learning controllers [22, 23].

### 2.2 Learning Models

In his visionary work, Kenneth Craik [24] hypothesized that if an organism carries a small-scale model of the world and its own possible actions in its head, it can use it to simulate different alternatives and decide on actions to take. This view has been further developed into mental models theory of cognition [25], theory of mind [26], and motor control [27].

Similarly, our robots could build mental models and use them to plan and accomplish tasks they encounter. Even if the task the robot faces is new, the world still works the same way and the same models apply. For example, the physics of the world do not change whether one is throwing a ball in a park or learning to play tennis. Many such models can be acquired or initialized from observations.

One line of work on learning world models, popular in machine learning and robotics, focused on prediction in pixel space [28, 29, 30, 31]. Representations obtained in this way can then be used for learning controllers [32]. While pixels are a nice and general way to obtain supervision, one may benefit from more abstract representations. Another line of work focused on learning using higher-level representations of entities and objects [33, 34, 35, 12].

When it comes to learning models of other agents, robotics and computer vision communities have focused on predicting future human trajectories in social navigation settings [36, 37, 38] as well as making more fine-grained predictions at the level of individual body parts [39].

### 2.3 Learning Behaviors

Studies of imitation in humans and primates have now spanned more than a century [40, 41]. Using Thorndike’s definition [40], imitation is “learning to do an act from seeing it done”. There is now a large body of work showing that humans learn to perform various behaviors through imitation.

Naturally, the idea of having robots learn to reproduce observed behaviors is an old one and it has been studied under a variety of names in robotics, including learning by watching, programming by demonstration, learning from demonstrations, and imitation learning [42, 43].

There are now many different problem settings and the specifics vary considerably. For example, the approach most commonly used in practice is direct behavior cloning [44]. This setting makes a number of assumption including access to teacher actions, same robot morphology, and dynamics.
We highlight the more general setting of learning from observations alone, and by learning behaviors we mean learning \textit{how} to perform various tasks from observations. Thanks to YouTube and computer vision datasets [45, 46, 47], we now have large quantities of observational data showing humans performing a wide variety of behaviors that we can learn to perform on a robot.

The learnt behaviors could be used as a whole or in a compositional fashion. For example, the robot could replicate human behavior based on latent representations [48, 49], or more structured representations of humans [50] and hands [51]. The robot could also learn only parts of observed behaviors and recombine them in new ways to achieve its goals [52, 53, 54].

In addition to learning specific human behaviors, the robot could leverage observations of humans in order to learn behaviors by itself. For example, the robot can see that a child plays with toys and that interesting parts of the space involve interacting with objects. Thus, observations can serve as a form of \textit{exploration} strategy for the robot and speed up learning and discovery of behaviors.

\section*{2.4 Learning Goals}

Developmental psychologists distinguish between imitation and emulation [55]. In particular, imitation can be defined as copying the \textit{form} of an action and emulation as copying the \textit{outcome} of an action sequence. We capture the notion of emulation into the ability to infer and understand goals. In other words, understand \textit{what} someone is trying to do by observing them.

In reinforcement learning terminology, this can be seen as learning the reward function. This has been explored in robotics [56], inverse reinforcement learning [57, 58], and reinforcement learning [59, 60, 61]. Similarly, this can be seen as a criterion, a cost function, a discriminator, or a critic. But really we want to capture a general notion of a goal and what it means to achieve it.

We can acquire such notions of goals from observational data. Consequently, these can be used to guide learning of behaviors through interactions, e.g. by serving as a reward functions for RL, or they could be coupled with a learnt models to evaluate actions and make decisions. For example, very much like Craik envisioned, the robot could use learnt models to determine possible outcomes of potential actions, rank them using a learnt notion of success, and pick the best action.

\section*{3 Discussion}

In this article, we have presented a view in favor of robot learning from observations and constructed a taxonomy of observational learning. In the following, we briefly answer some of the questions that might arise when discussing this article.

\textbf{What is new in this article?} People have, of course, already been working on different aspects of observational learning and we hope the text makes that clear. The goal of this article is to unify the related efforts under a common framework and bring community attention to this research direction.

\textbf{Is observational data enough?} No, it is not. While a lot can be learnt or accelerated by learning from observation alone, we still require learning through interaction. The two are complementary.

\textbf{Why treat observational data separately?} One could argue that a robot can always choose to observe if it wants to and that there is nothing special about it, it is just another action. And he would not be wrong, just not very practical. We believe that it is useful to distinguish between the two, to ensure that we develop technique capable of leveraging both forms of data effectively.

\textbf{Why does robot learning need to be inspired by humans?} It does not. Here we focus on a human-inspired approach as we believe that it is a promising path for developing general purpose robots that operate in human environments and that it may facilitate our understanding of humans.

\textbf{Do we need unsupervised, supervised, or reinforcement learning?} Learning from observational data likely involves at least a subset of them, if not all. And maybe more. Here we focus on the problem setting and its different aspects, rather than the classes of learning methods required for it.
References


