Robotics and Autonomous Systems 65 (2015) 24–39

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Applying Ant Colony Optimization algorithms for high-level behavior learning and reproduction from demonstrations

Benjamin Fonooni [a,*], Aleksandar Jevtić [b,c], Thomas Hellström [a], Lars-Erik Janlert [a]

[a] *Department of Computing Science, Umeå University, 901 87 Umeå, Sweden*

[b] *Robosoft S.A, Technopole Izarbel, 64210 Bidart, France*

[c] *École supérieure des technologies industrielles avancées (ESTIA), Technopole Izarbel, 64210 Bidart, France*

HIGHLIGHTS

- A behavior learning method based on Ant Colony Optimization (ACO) is proposed.
- We combined ACO, Semantic Networks and Spreading Activation mechanisms.
- The method is able to learn high-level aspects of behaviors from demonstrations.
- The method answers questions of "What to imitate" and "When to imitate".
- The method generalizes concepts while learning high-level aspects of behaviors.

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form
21 November 2014
Accepted 1 December 2014
Available online 10 December 2014

Keywords:
Learning from Demonstration
Semantic networks
Ant Colony Optimization
High-level behavior learning

ABSTRACT

In domains where robots carry out human's tasks, the ability to learn new behaviors easily and quickly plays an important role. Two major challenges with Learning from Demonstration (LfD) are to identify what information in a demonstrated behavior requires attention by the robot, and to generalize the learned behavior such that the robot is able to perform the same behavior in novel situations.

The main goal of this paper is to incorporate Ant Colony Optimization (ACO) algorithms into LfD in an approach that focuses on understanding tutor's intentions and learning conditions to exhibit a behavior. The proposed method combines ACO algorithms with semantic networks and spreading activation mechanism to reason and generalize the knowledge obtained through demonstrations. The approach also provides structures for behavior reproduction under new circumstances. Finally, applicability of the system in an object shape classification scenario is evaluated.

1. Introduction

During the past years robot task learning has received remarkable attention and motivated the robotics community to take a deeper interest in techniques based on human skill learning from observation [1]. In robotics, such an approach fits in the framework of Learning from Demonstration (LfD).

In the current paper, we address the questions "*What to imitate*" and "*When to imitate*" from a high-level perspective, while employing methods to learn and reproduce motor actions from demonstrations. Therefore, our focus is on learning and reproducing high-level aspects of demonstrated behaviors. For this purpose, we utilize a core Semantic Network (SN) as a model to represent behaviors by nodes and link them to a set of other nodes, which correspond to concepts and objects in the real world. The network contains concepts, objects and their properties required for learning and reproducing behaviors, and must be provided prior to the learning and reproducing process. The learned behaviors are then used as object affordances as well as preparing the ground for behavior arbitration. Our learning methods also provide techniques to learn conditions that result in the behavior and thus answer the question of "*When to imitate*". These conditions can be environmental, objects to use, and concepts related to demonstrated behavior. Therefore, depending on the amount of knowledge available in the core SN, the robot may perceive enormous amounts of information during learning. In many cases when the desired behavior is very complex or is demonstrated in an ambiguous manner, the robot requires a bias in order to focus on the right aspects of the demonstration [2]. By having a controller from a higher

* Corresponding author.
E-mail addresses: fonooni@cs.umu.se (B. Fonooni), a.jevtic@net.estia.fr (A. Jevtić), thomash@cs.umu.se (T. Hellström), lej@cs.umu.se (L.-E. Janlert).

http://dx.doi.org/10.1016/j.robot.2014.12.001

abstraction level to guide the robot especially during the learning phase, the robot's attention can be directed to aspects of demonstration that are significant for learning behaviors and thereby answering question of "*What to imitate*". This controller is a part of an architecture that has been proposed and developed in our previous work [3,4], and is employed in the current paper.

Our current work introduces, as an original contribution, the use of SNs for biasing the robot and the use of ACO algorithms to learn new behaviors and define the degree of generalization in order to exhibit the learned behaviors in situations new to the robot. Our previously developed learning method [4] is based on one-way ANOVA test and has some observed limitations caused by the imposed statistical constraints. This makes the method less successful in noisy environments. To address this issue, our new method views the problem of behavior learning as an optimization problem and applies ACO algorithms to determine the most related elements of demonstrations. Nodes and links in the core SN represent elements of a demonstration and the goal is to find the shortest path between each element's node and a behavior node using pheromone-laying behavior implementation of ACO.

The rest of the article is organized in the following manner: Section 2 provides background of the work, Section 3 presents principle elements of SNs for modeling the world and representing behaviors, Section 4 provides adapted formulations of ACO algorithms for the purpose of learning behaviors, Section 5 elaborates learning and reproduction of high-level representation of behaviors using ACO and SN, Section 6 presents the experimental setup and results, Section 7 gives a discussion of the approach and finally in Section 8 we make the conclusions.

2. Background

2.1. Learning from demonstration

LfD is a promising way to naturally and intuitively teach robots new behaviors (skills) by demonstrating how to perform the behavior [5]. Applying LfD does not require any expert knowledge of robotics or domain dynamics, so it can be easily applied by non-roboticist users for both trivial and non-trivial behaviors. To apply LfD, a number of questions have to be answered, as brought to attention by Schaal [6] and Demiris and Hayes [7]. These central questions are known as the "Big Five" and include "*Who to imitate?", "When to imitate?", "How to imitate?", "What to imitate*?" and "*How to evaluate successful imitation*?" [8]. An extensive overview of LfD and the "Big Five" can be found in [1,5,9]. Among the five questions, *How* and *What* are mostly studied in the literature, and approaches to face these challenges are both high- and low-level. The low-level approaches deal with mapping of sensory-motor information, which produce actions that are performed by the robot's actuators. Works by Mataric [10], Dillmann [11], Ekvall and Kragic, [12], Pastor et al. [13], Billing and Hellström [14] and many others address this low-level perspective of LfD. The high-level approaches focus on the tutor's intention, the goal of a demonstration and to which objects, concepts or environmental states the robot should direct its attention. In order to truly imitate observed behaviors, understanding goals and results of actions are necessary. Otherwise the robot may successfully reproduce demonstrated motor actions without achieving the desired goal or result. Therefore, developing a sophisticated attention mechanism to identify the most important elements of demonstrations is essential. Works by Mahmoodian et al. [15], Hajimirsadeghi et al. [16], Cakmak et al. [17], Erlhagen et al. [18] and Chao et al. [19] address these challenges of conceptualization and goal identification from demonstrations.

In the LfD learning process, the tutor first demonstrates a desired behavior to the robot that produces mappings of sensory-motor states and generalizes the mappings to other similar situations. There are mainly two ways of conducting a demonstration to a robot [20]:

1. *Direct learning*: In this approach the tutor demonstrates a behavior directly by manually steering the robot using a device such as a joystick.
2. *Indirect learning*: In this approach the tutor performs a behavior and the robot learns it by observation. Usually, vision and/or some other sensing methods are applied to record the demonstration. In the current work, indirect learning is applied, and RFID sensing is used to observe and record environmental states containing objects that appeared during demonstrations. This method is explained in detail in Section 5.

Generalization is not only applied to low-level skills, but also has a significant role in learning concepts and high-level representation of behaviors. According to Mitchell et al. [21], generalization can be defined as a process of identifying common features by observing a set of training examples and forming a concept definition based on these features. We are interested in designing learning methods that are able to extend the knowledge from learned behaviors under known circumstances, to novel situations. Thus, the proposed learning method generalizes the high-level aspects of behaviors in order to reproduce them later even if the conditions are slightly changed.

2.2. Ant colony optimization

ACO is a metaheuristic that has been used to solve numerous complex engineering problems that can be represented as discrete optimization problems [22]. ACO implements the pheromone-laying behavior that natural ant colonies use to store the information about the environment, which can then be locally accessed by any member of the colony. In most cases, the goal of applying ACO is to find the shortest path between the points in the solution space, or to extract the accumulated pheromone pattern [23].

The first successful ACO algorithm was introduced by Marco Dorigo who was inspired by biological works of Deneubourg and colleagues [24]. They proposed a stochastic model of ant behavior by observing ant colonies and how they are searching for the shortest path between food sources and their nest [25]. The algorithm simulates foraging behavior of Argentine ants, which is to explore the surrounding area randomly and leave a pheromone trail on the ground while moving. In case of finding food, on their way back to the nest, they will leave a trail of pheromone whose quantity depends on the quality of the found food. This will guide other ants to choose the path that leads to high quality food by tracing the paths with strong pheromone concentrations [26]. This way of indirect communication between the ants is named *stigmergy* [27,28]. The principle of stigmergy is also applied to artificial ants in the ACO and the way they explore a discrete n-dimensional solution space [29]. However, the pheromone accumulation only provides a positive feedback mechanism to enhance the found solutions; in order to avoid pheromone saturation and allow the ants to always search for better solutions, negative feedback through pheromone evaporation is applied, that restrains the ants from exclusively taking the initially found path.

The general design methodology for Swarm Intelligence tools proposed in [23] is used to develop the ACO-based learning algorithms. The proposed methodology is given in Section 5.1.1.

2.3. Semantic networks and behavior representation

The main concern in modeling the world is how to structure and generalize information. Semantic Networks (SNs) are common techniques to represent abstract knowledge in systems. In robotics, SN is used more often for concept forming, situational awareness [30] and task planning [31]. In concept forming, monitoring the robot's attentional state to understand the amount of