Future Generation Computer Systems 64 (2016) 78–87

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Principles and experimentations of self-organizing embedded agents allowing learning from demonstration in ambient robotics

Nicolas Verstaevel[a,b,*], Christine Régis[a], Marie-Pierre Gleizes[a], Fabrice Robert[b]

[a] IRIT, Université Paul Sabatier, 118 rte de Narbonne, 31062 Toulouse Cedex, France
[b] Sogeti High Tech, 3 Chemin Laporte, 31300 Toulouse, France

HIGHLIGHTS

- Extreme Sensitive Robotic as a bottom-up approach to deal with complexity in ambient robotic.
- Self-adaptive multi-agent system for learning from demonstration in ambient robotic.
- Experiments show that tutorship learning with Context agent is promising.

ARTICLE INFO

Article history:
Received 30 September 2015
Received in revised form
25 February 2016
Accepted 28 March 2016
Available online 5 April 2016

Keywords:
Adaptive agents
Cooperative multi-agent systems
Robotics in ambient systems
Machine learning
Context-awareness

ABSTRACT

Ambient systems are populated by many heterogeneous devices to provide adequate services to their users. The adaptation of an ambient system to the specific needs of its users is a challenging task. Because human–system interaction has to be as natural as possible, we propose an approach based on Learning from Demonstration (LfD). LfD is an interesting approach to generalize what has been observed during the demonstration to similar situations. However, using LfD in ambient systems needs adaptivity of the learning technique. We present ALEX, a multi-agent system able to dynamically learn and reuse contexts from demonstrations performed by a tutor. The results of the experiments performed on both a real and a virtual robot show interesting properties of our technology for ambient applications.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Once confined in a science of control in structured environments, researches on robotics are now considering the integration of robotic devices in the real world for applications where tasks are diverse, complex and evolutive [1]. Service robotic differs from its industrial version by the interest in providing services to humans. Consequently, research now consider the use of robotic devices in ambient applications [2] and the term *ambient robotics* directly refers to the usage of robotic components in ambient systems. Ambient systems are characterized by their high dynamic and their complexity. Many heterogeneous robotic devices can appear and disappear along the system life-cycle and interact opportunistically together. According to the definition of Russell and Norvig [3], the environments of ambient systems are:

- **Inaccessible**: each device composing the system has a partial observation of the environment.
- **Continuous**: considering applications in the real world, the number of observations and actions is not discrete.
- **Non-deterministic**: consequences of performed actions in the real world could not be determined in advance with certainty.
- **Dynamic**: system's actions, user activity, appearance and disappearance of devices may change the environment.

Consequently, designing an *ad hoc* controller of a robotic device in an ambient system is a complex task that requires a lot of knowledge. This complexity is increased if we take into account that users have multiple, specific and often changing needs. Providing to those devices the ability to learn and adapt to users' needs is then a particularly challenging task [4]. To be as natural as possible, such learning ability needs to rest on a process that does not require any kind of technical knowledge for users (i). Furthermore, it needs both genericity, to be applicable on any kind of devices with any kind of users, and openness properties to deal with the appearance and disappearance of devices. The genericity and openness properties require then using agnostic learning techniques that makes as

* Corresponding author at: IRIT, Université Paul Sabatier, 118 rte de Narbonne, 31062 Toulouse Cedex, France.
E-mail address: nicolas.verstaevel@irit.fr (N. Verstaevel).

http://dx.doi.org/10.1016/j.future.2016.03.023
0167-739X/© 2016 Elsevier B.V. All rights reserved.

few assumptions as possible [5] (ii). To deal with (i) and (ii), we propose to use Learning from Demonstration (LfD), a paradigm to dynamically learn new behaviours. The paper is organized as follows: first, we present the problems and challenges of LfD in ambient systems. Then, we present ALEX our solution to handle this challenge. Two experiments are then proposed to illustrate ALEX's behaviour. At last, the conclusion discusses perspectives and future works.

2. Learning from demonstration

2.1. General principle

Learning from Demonstration, also named "Imitation Learning" or "Programming by Demonstration", is a paradigm mainly studied in the robotic field that allows systems to self-discover new behaviours [6]. It takes inspiration from the natural tendency of some animal species and humans to learn from the imitation of their congeners. The main idea is that an appropriate controller for a robotic device can be learnt from the observation of the performance of another entity (virtual or human) named as the *tutor*. The tutor can interact with the system to explicit the desired behaviour through the natural process of demonstration. A demonstration is then a set of successive actions performed by the tutor in a particular context. The learning system has to produce a mapping function correlating observations of the environment and tutor's actions to its own actions. The main advantage of such technique is that it needs no explicit programming or knowledge on the system. It only observes tutor's actions and current system context to learn a control policy and can be used by end-users without technical skills.

The paradigm has been used on a wide range of applications such as autonomous car following [7], robot trajectory learning [8], robot navigation in complex unstructured terrain [9] or haptic guidance of bimanual surgical tasks [10]. Recent surveys [11,6] propose an overview of the LfD field illustrating a wide variety of applications. Our interest is not to focus on one particular application. On the contrary, we want to deal with any kind of ambient robotic system. This section ends with a study of the usage of the LfD paradigm in the context of ambient robotics.

2.2. Problem formalization

Billing and Hellström [12] propose a complete LfD formalization. Here, we propose to introduce the fundamental concepts of LfD and discuss its usage in ambient robotics.

LfD is a problem of imitation as an entity tries to produce a behaviour similar to another entity. A tutor evolving in a world realizes an observation Ω of this world. The tutor can perform a set of actions A (A could be empty) and follows a policy π_{tutor} (1) that associates to any world state a particular action. It is supposed that (1) is the optimal policy to satisfy the tutor.

- $\pi_{tutor} : \Omega \to a \in A.$ (1)

The learner disposes of a set of observations O (named observation space) on the world and its own set of actions B. The learner follows another policy $\pi_{learner}$ (2) in order to produce a behaviour similar to the observed one.

- $\pi_{learner} : O \to b \in B, \; \pi_{learner} \equiv \pi_{tutor}.$ (2)

In most cases, the tutor and the system, while evolving in the same world (which can be a virtual world or the real world), have a different observation of the world. It is particularly true in real world problems with human tutors where the world is observed by the system through sensors whereas the human observes it through its own senses. It results in a problem of perception equivalence [13]. The tutor demonstrating a particular behaviour can observe modifications of the world that the learner cannot perceive. However, equivalences of perception can be found by the system. For example, a user is cold and turns the heating on. Observing through sensors that a user is cold is complex, but observing the current temperature, wind and humidity levels is easily feasible. A learner can make a correlation between the current situation described by sensors and the action of turning the heating on and it can learn that it is necessary to turn the heating on when a similar situation occurs. The learner has to find correlations between its own observations and the performance of an action by the tutor. This raises the challenging question of the possible lack of perception. A tutor can perform a demonstration dependent of a phenomenon that the learner cannot observe. In this paper, we consider that the learner has sufficient observations to perform the task. Nevertheless, some clues to handle this problem are proposed in Section 5 as perspectives. The problem to tackle is then how to interpret those observations to construct a control policy enabling the learner to produce an imitative behaviour.

2.3. Lfd and ambient systems

Ambient systems are rich of interaction possibilities for users. We propose to exploit the inherent interactivity of ambient systems in order to learn and adapt from users' activity. LfD then appears to be a suitable paradigm to learn and generalize recurrent activities from the observation of users' activity. Therefore, users can be seen as system's tutors. Each user's action on a device is then seen as a demonstration of the desired behaviour. The key idea is if a user has to act on devices, the reason is that the user wants to change his environment to improve the current service. A consequence is that user actions are minimized and so functionalities are more relevant. The device can then use this information to self-adapt. However, ambient systems have some particular properties and require functional properties that are challenging for LfD.

Five central questions have been identified that need to be addressed by scientists interested in designing experiments on imitation [14,15]:

- **Who** to imitate: first is the choice of the model (the one to be imitated).
- **When** to imitate: second is determining when imitation has to be done. There are typically two types in literature if whether the imitation is immediately leading to synchronous behaviour or deferred which means that the imitated behaviour might occur even in the absence of the model.
- **What** to imitate: a distinction has to be made between goals, actions and results and which part of the demonstration has to be replicated.
- **How** to imitate: the ***how*** question addresses the problematic of generating an appropriate mapping between the model behaviour and the imitator's one.
- What is a **successful** imitation: one needs to be able to distinguish good imitations from bad imitations. Then, good metrics must be clearly identified.

The application of LfD in the robotic field mainly focuses the "What" and "How" questions [6]. On a previous paper, we position ourselves regarding those questions [16] and discuss specificities of ambient systems.

An ambient system is open, which means that entities can appear and disappear during system activity. This openness property is challenging for LfD, as it does not allow doing assumptions on system's composition. With respect to the openness property, it must be considered that the set of observations O is a priori unknown. One must deal with situations of opulence of data (which induces that some data may be useless in order to learn the task)

Download English Version:

https://daneshyari.com/en/article/424786

Download Persian Version:

https://daneshyari.com/article/424786

Daneshyari.com