Teaching Human-like Movements to Autonomous Agents via Implicit Training

Anton Bogdanovych[1], Marc Esteva[2], Simeon Simoff[1] and Carles Sierra[2]

[1] Faculty of Information Technology,
University of Technology Sydney
2007 NSW, Australia
{anton, simeon}@it.uts.edu.au
[2] Artificial Intelligence Research Institute (IIIA-CSIC)
Campus UAB, Barcelona, Catalonia, Spain
{esteva, sierra}@iiia.csic.es

Abstract. This paper provides a brief overview of an implicit training method used for teaching autonomous agents to represent humans in 3D Virtual Worlds without any explicit training efforts being required.

1 Introduction

Many scholars, whose work is focused on intelligent virtual agents face the problem of making them believable. The believability has a lot of different characteristics, e.g. personality, social role awareness etc [1]. In our work, instead of trying to discover and implement different characteristics of believability, we use imitation learning [2]. The main hypothesis behind it can be best summarized by the cliche "to know a man is to walk a mile in his shoes" [2]. To increase the believability of agents we suggest that they constantly observe the behavior of humans and learn to imitate it. Efficient imitation can be achieved when a human is fully immersed into an environment based on the 3D Electronic Institutions technology [3]. This technology utilizes 3D Virtual Worlds, used for the visualization purposes, together with Electronic Institutions that help to establish the rules of the interactions among participants. Such approach valuably facilitates the training of autonomous agents. 3D representation of the environment provides as much possibilities to observe the behavior of the humans as the real world does. It assumes similar embodiment for all participants, including humans and autonomous agents who imitate the humans, so every action that a human performs can be easily observed and then reproduced by an autonomous agent, without a need to overcome the embodiment dissimilarities. Moreover, the use of Electronic Institutions provides context and background knowledge for learning, helping to explain the tactical behavior and goals of the humans.

2 Implicit Training

An important feature of 3D Electronic Institutions is that every human participant is always supplied with a corresponding software agent that communicates the desires of the human to the institutional infrastructure. While a human drives an avatar and acts in the Virtual World, the agent observes those actions and learns how to make the decisions on human's behalf. 3D Electronic Institutions

separate the actions that happen in the Virtual World into two different kinds: normative level actions (operations that require institutional validation) and visual level level actions (no validation required). An example of a normative level action is opening a door to enter a secure auction. An example of a visual level action can be a gesture or any other kind of avatar movement.

Based on these actions, the learning-related information for each of the agents is stored in a separate learning graph. The nodes of this graph correspond to normative level actions. Each node is associated with two variables: the action name together with parameters and the probability ($P(Node)$) of this action to be executed. The arcs connecting the nodes are associated with prerecorded sequences of visual level actions $(s_1, \ldots, s_n)$ and the attribute vectors that influenced them $(a_1, \ldots, a_n)$. Each pair $\langle a_n, s_n \rangle$ is stored in a hashtable, where a_i is the key of the table and s_i is the value. Each a_i consists of the list of parameters: $a_i = \langle p_1, \ldots p_k \rangle$. We assume that the behaviour of the principle is only influenced by what is currently visible through the field of view of the avatar and limit the visible items to the objects located in the environments and other avatars. So, the parameters that can be used for learning are recorded in the following form: $p_i = \langle V_o, V_{av} \rangle$. Here V_o is the list of visible objects; V_{av} is the list of visible avatars. Each time an institutional message is executed, the autonomous agent records the parameters it is currently able to observe, creates a new visual level sequence and every 50 Ms adds a new visual level message into it. The recording is stopped once a new institutional message is executed.

The nodes of the learning graph are seen as internal states of the agent, the arcs determine the mechanism of switching from one state to another and the probability $P(Node)$ determines how likely it is for the agent to change its current state to the state determined by the next node. Once the agent reaches a state $S(Node_i)$ it checks all the other nodes connected to the $Node_i$, selects the node ($Node_k$) with the highest probability and changes its current state to $S(Node_k)$ by executing the best matching sequence of the visual level actions recorded on the arc that connects $Node_i$ and $Node_k$. The best matched sequence is selected through the employment of a classifier, which compares the currently observed parameters to the parameters associated to the recorded sequences.

3 Conclusion

We presented the implicit training method and described an approach to implement it using a learning graph. Future work includes further development of this concept and its evaluation.

References

1. Magnenat-Thalmann, N., Kim, H., Egges, A., Garchery, S.: Believability and interaction in virtual worlds. In: MMM. (2005) 2–9
2. Breazeal, C.: Imitation as social exchange between humans and robots. In: AISB Symposium on Imitation in Animals and Artifacts. (1999) 96–104
3. Bogdanovych, A., Berger, H., Simoff, S., Sierra, C.: E-Commerce Environments as 3D Electronic Institutions. In: Proc. of IADIS e-Commerce 2004, Portugal (2004)