

LEARNING HIGH-LEVEL BEHAVIORS FROM DEMONSTRATION THROUGH SEMANTIC NETWORKS

Benjamin Fonooni, Thomas Hellström and Lars-Erik Janlert
Department of Computing Science, Umeå University, SE 901 87, Umeå, Sweden

Keywords: Learning from Demonstration, High-Level Behaviors, Semantic Networks, Robot Learning.

Abstract: In this paper we present an approach for high-level behavior recognition and selection integrated with a low-level controller to help the robot to learn new skills from demonstrations. By means of Semantic Network as the core of the method, the robot gains the ability to model the world with concepts and relate them to low-level sensory-motor states. We also show how the generalization ability of Semantic Networks can be used to extend learned skills to new situations.

1 INTRODUCTION

Learning from Demonstration (LfD) is a technique to teach robots new behaviors by having a human or robot teacher performing sequences of actions that are either observed or perceived by the robot. Several algorithms have been proposed. Most of them distinguish between low and high-level representations of a behavior (see for instance Billard et al., 2008). In our approach, the low-level is represented by sensory-motor mappings and the high-level by combinations of concepts represented in Semantic Networks.

One of the hard problems in LfD is how to generalize a demonstrated behavior such that it can be performed also in new, previously unseen situations. This issue exists from both high and low-level perspectives and there are several ways to approach it (Nehaniv and Dautenhahn, 2000; Byrne and Russon, 1998). The purpose of this paper is to introduce a technique that integrates high and low-level learning and control in a way that supports generalization. A high-level controller deals with concepts represented and processed in Semantic Networks (SN). This controller is interfaced to a low-level controller that learns and performs behaviors defined at the sensory-motor level. The glue, interfacing the two levels, is learned *contexts*, describing the necessary high-level conditions for a low-level behavior to be performed.

Behavioral studies of animals and humans provide several sources for ideas on how low and

high-level learning may be combined. For instance, *Goal emulation* (Whiten and Ham, 1992) is interesting for learning how to direct focus of attention towards favorable goals. *Stimulus enhancement* is the implicit memory effect when stimuli in the environment bias the agent's behavior towards receiving similar stimuli in the future. *Response facilitation* is a mechanism that describes how motor responses already in the repertoire be repeated after observing the performance of the same action (Kopp and Graeser, 2006). In a broad sense, the work presented in this paper may be seen as a realization of response facilitation. All mechanisms described above may be seen as examples of priming, aiming at guiding animal behavior and learning (Byrne, 1994).

In section 2, the proposed architecture with its major units is described. Section 3 is an overview on Semantic Networks and its features. Section 4 elaborates the learning and performing phases based on proposed architecture and Semantic Networks. In section 5, an example is shown to evaluate learning and performing phases.

2 OVERVIEW OF THE ARCHITECTURE

A number of architectures and frameworks for LfD have been developed during passed years which influenced the current research in this field (Kasper et al., 2001; Nicolescu, 2003). These works are

introducing architectures for learning low-level sensory-motor behaviors. The purpose of designing a new architecture is to interface the low-level behavior learning and control which introduces an integration and behavior arbitration techniques by means of high-level control. Figure 1 depicts the proposed architecture. The major units are described below.

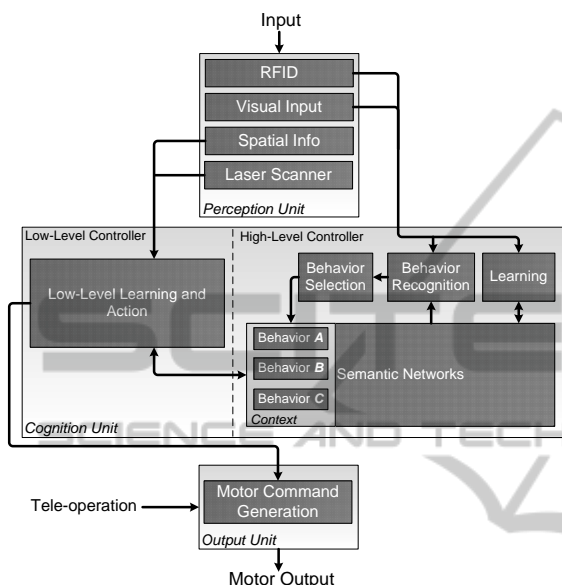


Figure 1: Proposed architecture.

2.1 Perception Unit

This unit collects and pre-processes sensor data from the environment. In our experimental set-up, the robot is equipped with laser scanner, ultrasonic transducers, infra red sensors and an RFID reader that acts as a high-level sensor which delivers identity (and to some extent position) of places, objects and people equipped with RFID tags. Every tag is associated with a number of properties defined in a database. The RFID technique is commonly used in robotics to get reliable performance and human-robot interaction (Nguyen et al., 2009), and should be considered as a complement to other high-level sensors like face recognizers, emotion detectors, gesture recognizers or any other visual inputs, and not a replacement.

2.2 Cognition Unit

The Cognition Unit is responsible for all robot decision making and action selection processes. It contains representation of the robot's cognitive state and has functions to modify its internal states.

Due to the structural differences between low and high-level information, the unit is organized in two modules running simultaneously.

2.2.1 High-Level Controller

One of the main tasks for this module is to learn contexts that are relevant for the execution of low-level behaviors. The other task is to arbitrate the low-level controller. The module relies on the abilities of a Semantic Network with predefined concepts and relations describing the environment. In learning mode, high-level inputs from the Perception unit update the SN such that contexts associated with the demonstrated behaviors are learned. In execution mode (performing phase), the module supports the low-level controller by activating the most relevant context(s) according to the current environmental conditions. These contexts act as bias in the activation of behaviors in the low-level controller. Basically, the cognitive state of the robot, represented in the Semantic Network, is updated through perception and a behavior recognition process, and acts as a cue for performing a behavior.

2.2.2 Low-Level Controller

The low-level controller learns and executes behaviors that are mappings of sensory-motor data to low-level actions (Billing et al., 2010a; Billing and Hellström, 2010b). In the presented work, the technique for learning is *Predictive Sequential Learning (PSL)* (Billing and Hellström, 2008). PSL treats control as a prediction problem and decides the next action based on the sequence on recent sensory-motor events. This technique allows learning of many types of complex behaviors, but does only work as long as the recent sensory-motor history contains all information necessary to select an appropriate action. One way to overcome this limitation is to define several contexts for PSL, where each context acts as a bias for action selection. In this way, actions that are less common within the present context are inhibited and the risk for selecting inappropriate actions is reduced.

2.3 Output Unit

This unit is designed to enable tele-operation of the robot. In addition, execution of action selection results coming from the Cognition unit into motor commands will be performed here.

3 SEMANTIC NETWORKS

Semantic Networks are often used to represent abstract knowledge in a human-like fashion. They are common within artificial intelligence as well as in philosophy, psychology and linguistics (Bagchi et al., 2000; Brown and Cocking, 2000; Rodriguez, 2008). In robotics, Semantic Networks is used for concept forming and situational awareness (Coradeschi and Saffiotti, 2003). The structured way of representing knowledge can in combination with visualization tools (Hartley and Barnden, 1997) help humans to understand the internal state of the robot and what is happening in the robot's cognitive system. This may for instance help a tutor to put the robot back on track when it is distracted during learning or performing phases.

In our usage of Semantic Networks, high-level concepts such as perceived object types and properties are represented as nodes while relations between concepts are represented as links. The initial SN is pre-defined and comprises nodes that are connected to the perception unit. These nodes are activated through perception.

3.1 Generalization Ability

A common reason for using a Semantic Network as a model of the environment is its ability to generalize (Mugnier, 1995), (Vashchenko, 1977). In our case, after a demonstration in LfD, the robot will be able to extend the learned context to other, related, contexts. Assume for instance that the robot learns how to clean the table if there are empty cups on it. By generalizing the cup concept to all the drinkwares, it will also perform the cleaning behavior when perceiving a mug on the table.

3.2 Interfacing to Low-level Information

The success of robots designed to learn and work in daily environments with humans, relies on wrapping sensory-motor information with high-level concepts. This can improve human-robot interaction by utilizing Semantic Networks (Galindo et al., 2005). As mentioned earlier in section 2.2.2, *contexts* which are activated by the Semantic Network, give meaning to low-level information and act as a bias to choose suitable behaviors.

3.3 Spreading and Decaying Activation

In the proposed approach, each node has an

activation level. *Spreading* is a mechanism by which activation spreads from one node to another in proportion to the strength of their connection. *Decaying* is a mechanism by which the activation levels of nodes gradually decrease over time. These processes have been implemented in a variety of ways to solve different problems in modeling, learning and robotics (Bagchi et al., 2000; Brown and Cocking, 2000). The spreading activation model used in this work, is based on mechanisms of human memory operations that originates from psychological studies (Rumelhart and Norman, 1983) and was first introduced in computing science in the area of artificial intelligence to provide a processing framework for Semantic Networks (Crestani, 1997). In order to make spreading activation work properly, we made following assumptions:

- Activation spreads in parallel, to all links leading out from a node
- Activation at a node is divided among the links that lead from it
- Activation decays rapidly without stimulation from other nodes or inputs
- Each node has an energy parameter that limits the number of link levels for spreading

The degree of generalization depends on the amount of energy available for propagation of activations. Higher amounts allow spreading along several links, which leads to higher connectivity of nodes and increase generalization.

The connections between nodes have weight values that limit the propagation of activation through the network. Learning is used to modify the connection weights and will be discussed in the next section.

4 LEARNING AND PERFORMING PHASES

One of the objectives of the research presented in this paper is to develop mechanisms to identify high-level contexts in a demonstration, and map each context to a low-level behavior. The low-level controller is assumed to contain learning capability based on sensory-motor data, and an ability to execute the behaviors on request. In section 4.1 we describe how high-level contexts are learned simultaneously with the low-level learning.

4.1 Learning Phase

Our learning approach is inspired by Novelty Detection techniques which are commonly used to detect new situations that did not occur during training (Markou and Singh, 2003).

We assume that we already have a predefined Semantic Network based on an ontology of the domain in which the robot should operate. This network is used to interface to the Perception unit and to identify or activate related nodes through spreading and decaying activation.

The learning process starts by generating a *History Network* describing the normal state of the world. The environment is observed by sampling the high-level sensors at a given frequency. As mentioned earlier, RFID tags are used for simplifying object detection and identification. Each readout gives object identities and properties that are perceived in the environment. Each read tag causes the corresponding nodes to be activated. For instance, if the RFID belonging to a red ball is detected, the nodes *Red* and *Ball* will be activated. Throughout the learning process, activation levels propagate to all connected nodes by spreading activation.

Sometimes a node is activated and deactivated due to noise and uncertainties in the RFID sensing. Therefore, a decaying delay parameter is defined to prevent instant deactivation of nodes after the disappearance of correspondent object from the environment.

Finally, the updated Semantic Network will be saved as the *History Network*.

Now learning of a high-level context may start. A *context node* with the name of the new behavior to be learned is added to the network. This version of the network is called *Learning Network*. The human teacher then demonstrates the wanted behavior by tele-operation. The RFID sensors perceive high-level concepts, at the same time as sensory-motor data is learned by the low-level controller. The context node will be connected to nodes activated by the RFID sensors. To finalize the learning process, two issues must be solved. First, the most relevant connections must be determined. Secondly, the weights between the remaining nodes and the context node must be computed. In order to identify relevant connections, the algorithm looks for significant differences between the *History* and *Learning Networks*. An unpaired T-Test is used to compare mean node activation for all nodes.

$$t_x = \frac{\mu_{AHx} - \mu_{ALx}}{\sqrt{\frac{Var_{Hx}}{n_H} + \frac{Var_{Lx}}{n_L}}} \quad (1)$$

where

μ_{AHx} is mean activation of History Node x

μ_{ALx} is mean activation of Learning Node x

n_H and n_L are number of samples for History and Learning respectively

t_x tells whether the samples for the two nodes are drawn from the same distribution or not. In other words: did the node change significantly between History and Learning phases. If it did not, the connection between the node and the context node should be removed. For instance, suppose *ambient light* was always *on*, during both *History* and *Learning* phases. In this case, the T-Test will consider *ambient light* as irrelevant because of the identical distribution in both phases.

After the elimination process of irrelevant connections, weights (w_x) for the remaining nodes are calculated. This is done by the following equation:

$$w_x = \frac{N_x \mu_{ALx}}{P} \quad (2)$$

where N_x is the number of samples for which node x has activation value above 0 during the learning phase, and P is the weighted sum for all nodes, calculated as follows:

$$P = \sum_{i=1}^n N_i \mu_{ALi} \quad (3)$$

Finally, the learned context node will be associated with the learned behavior representation in the low-level controller module.

4.2 Performing Phase

In the performing phase, RFID sensors update their corresponding network nodes. Whenever a node is activated, all other linked nodes are activated according to the spreading mechanism. In this way previously learned context nodes may get activated, thereby, guiding the low-level controller to execute the behaviors. If two or more contexts have high activation levels, several behaviors are possible, and the final decision will be made by the low-level controller. This can be viewed as high-level behavior recognition and is performed by *Behavior Recognition* module depicted in Figure 1. Due to the pre-defined semantic relations in the semantic network, the robot will be able to generalize the demonstrated context to similar contexts. As

previously mentioned, the degree of generalization can be controlled by the amount of energy (Huang et al., 2006).

5 EXPERIMENTS

For better understanding of the whole approach, an example is shown. Assume we are going to teach the robot how to move a thing to the storage room. First, the robot will start moving around by tele-operation and collecting information regarding all the objects and places by RFID tags. Due to the characteristics of the described technique, the blue box should not be present at this stage. Figure 2 depicts the robot's perceptions that yield the *History Network*.

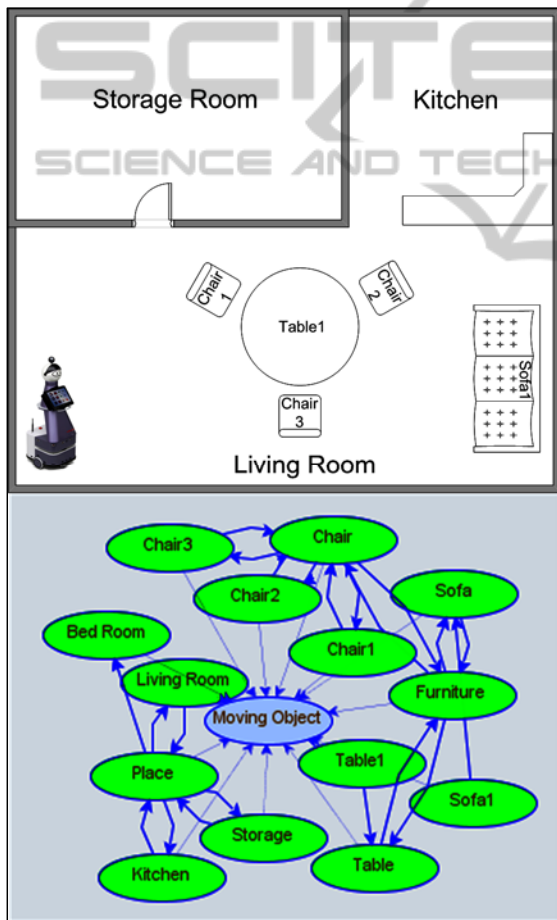


Figure 2: History network.

Learning will begin by placing the blue box somewhere in the middle of the room and tele-operating the robot towards the box. After grasping, the teacher commands the robot to push the box and guides it to the storage room that ends the learning

phase. Figure 3 depicts the learned *Moving Object* behavior. Although we did not illustrate any low-level learning, this is done simultaneously by the low-level controller while tele-operating the robot. The number of samplings for the history (N_H) and learning (N_L) is 40 and 60 respectively. In order to identify the nodes with the most significant changes, the t-test is run and results are shown in Table 1. Confidence Interval (CI) of the test is given by the t-distribution with α value set to 0.05. Degree of Freedom (DF) is calculated as follows:

$$DF = (N_H + N_L) - 2 \tag{4}$$

According to equation 1, t_x will be computed and nodes which fulfill condition 5 will be removed.

$$-CI \leq t_x \leq CI \tag{5}$$

Finally, according to equation 2, weights for the remaining nodes are calculated, shown in Table 1. After finalizing the learning phase, the robot is able to perform *Moving Object* action whenever it perceives *blue* and *box1* in the environment.

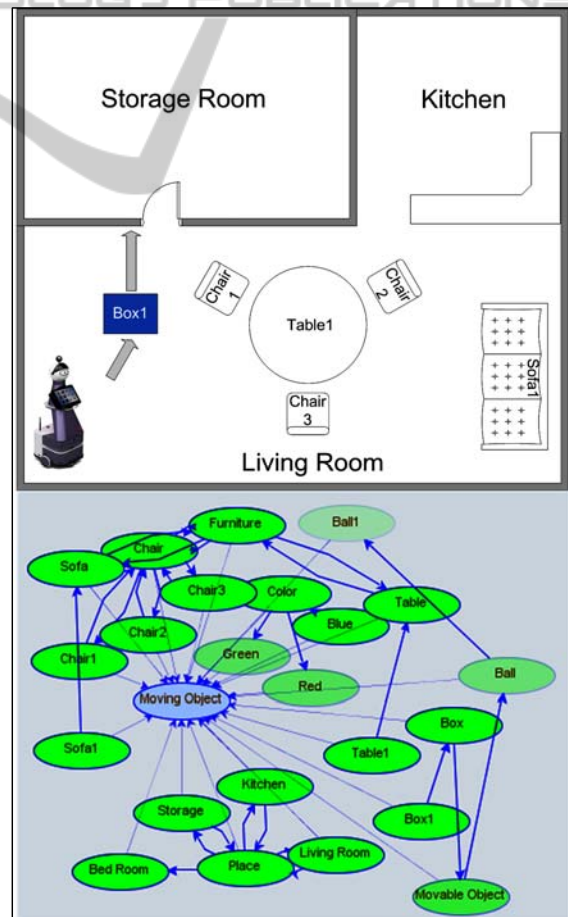


Figure 3: Learned network.

Table 1: History and learning values, T-Test results and weight values.

Node	μ_{AHx}	σ_{Hx}	N_{Hx}	μ_{ALx}	σ_{Lx}	N_{Lx}	t_x	DF	CI	P_x	w_x
Living Room	1.51	0.56	36	1.65	0.37	58	-0.48	98	2.0	0%	--
Sofa	1.82	1.2	28	2.24	0.94	51	-1.0	98	2.0	0%	--
Table1	1.73	1.16	28	2.15	0.9	51	-1.03	98	2.0	0%	--
Furniture	2.64	1.62	31	3.21	1.24	54	-0.96	98	2.0	0%	--
Box	0.0	0.0	0	0.73	0.44	44	-6.52	58	2.02	19.74%	0.1974
Box1	0.0	0.0	0	0.73	0.44	44	-6.52	58	2.02	19.74%	0.1974
Bed Room	1.3	0.48	36	1.41	0.32	58	-0.48	98	2.0	0%	--
Chair2	2.81	1.71	31	3.41	1.31	54	-0.95	98	2.0	0%	--
Table	2.12	1.40	28	2.61	1.09	51	-1.0	98	2.0	0%	--
Blue	0.0	0.0	0	0.66	0.47	40	-6.21	58	2.02	16.32%	0.1632
Chair1	2.26	1.40	31	2.76	1.07	54	-0.97	98	2.0	0%	--
Kitchen	0.87	0.31	36	0.95	0.20	58	-0.46	98	2.0	0%	--
Movable Obj.	0.0	0.0	0	0.36	0.22	44	-6.52	58	2.02	9.87%	0.0987
Green	0.0	0.0	0	0.21	0.15	39	-6.14	58	2.02	5.17%	0.0517
Chair3	3.21	1.89	31	3.87	1.43	54	-0.93	98	2.0	0%	--
Sofa1	0.67	0.46	27	0.85	0.35	51	-1.09	98	2.0	0%	--
Ball1	0.0	0.0	0	0.09	0.05	44	-6.52	58	2.02	2.47%	0.0247
Red	0.0	0.0	0	0.22	0.15	40	-6.21	58	2.02	5.44%	0.0544
Storage Room	1.97	0.69	36	2.14	0.43	58	-0.46	98	2.0	0%	--
Ball	0.0	0.0	0	0.18	0.11	44	-6.52	58	2.02	4.94%	0.0494
Place	2.62	0.94	36	2.85	0.60	58	-0.46	98	2.0	0%	--
Chair	3.65	2.13	31	4.39	1.61	54	-0.92	98	2.0	0%	--
Color	0.0	0.0	0	0.66	0.47	40	-6.21	58	2.02	16.32%	0.1632

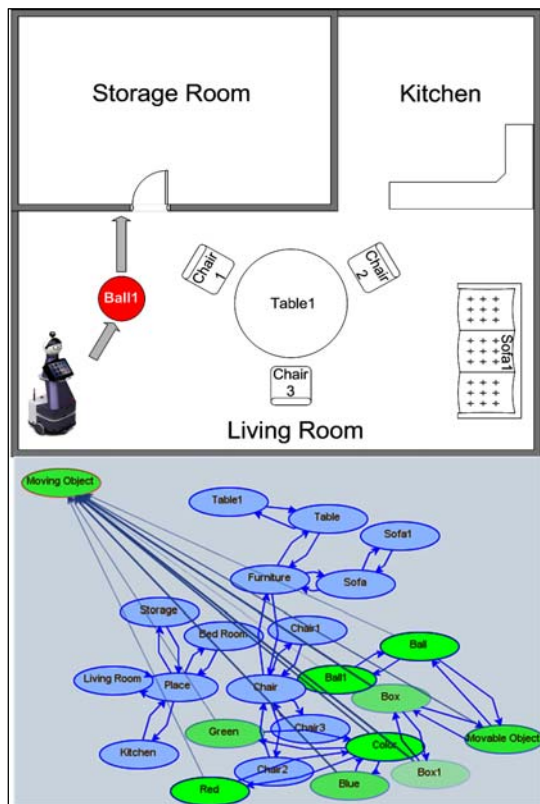


Figure 4: Performing phase.

The robot is not only capable of performing *Moving Object* behavior by observing the same objects during the learning phase, but can also generalize objects and concepts in the new situations. In order to test the system, the robot should recognize a *red ball* and push it to the storage room. This example clearly shows the generalization ability mentioned in section 3.1. As can be seen in Figure 4, perceiving *red* and *ball1* instead of *blue* and *box1*, to some degree, activates *Moving Object* context node through direct links and other connections to the *Color* and *Movable Object* nodes.

The activation level of the context node (A_c) is calculated by equation 6:

$$A_c = \sum_{i=1}^n A_i w_i \quad (6)$$

n is the number of nodes which are currently activated and connected to the context node. A selection threshold should be defined for accepting the selected behavior as a result of generalization. In our example, we set the threshold to 0.6 meaning that the result of equation 6 should be at least 60% of the maximum value of the context node's activation (A_{cmax}). The maximum value is calculated during the learning phase by equation 6 and by replacing A_i with A_{imax} (maximum

activation of node i). For this example, A_{cmax} equals 0.8214 and calculated A_c is 0.5246 which passed our threshold with 63%. Therefore, *red* and *ball1* are also able to trigger *Moving Object* context and cause the low-level controller to execute corresponding sensory-motor commands.

6 CONCLUSION AND FUTURE WORKS

In this paper we proposed an architecture to learn and act at a conceptual level by means of Semantic Networks. By introducing Semantic Networks and their usage in some research projects, a possible integration to LfD discussed. These aspects are valuable in concept forming and provide support for higher level cognitive activities such as behavior recognition. This integration is useful not only for LfD, but can be utilized in scaffolding, reinforcement learning or any other supervised learning algorithms. In this work, functionality of the system is tested with limited objects in the environment. In case of scaling up the number of entities in the working ontology, generalization will be more applicable.

Currently, our approach is incapable of handling quantities and negations. In our future work, we are going to define new link types in the Semantic Networks and design the high-level control in a way that can learn more complex scenarios.

ACKNOWLEDGEMENTS

This work has been financed by the EU funded Initial Training Network (ITN) in the Marie-Curie People Programme (FP7): INTRO (INTeractive Robotics research network), grant agreement no.: 238486.

REFERENCES

- Bagchi, S., Biswas, G., and Kawamura, K., 2000. Task Planning under Uncertainty using a Spreading Activation Network, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 30, No. 6, pp. 639-650.
- Billard, A., Calinon, S., Dillmann, R., and Schaal, S., 2008. Robot programming by demonstration. In Siciliano, B. and Khatib, O., editors, *Handbook of Robotics*. Springer.
- Billing, E. A., Hellström, T., and Janlert, L. E., 2010a. Behavior Recognition for Learning from Demonstration, *Proceedings of IEEE International Conference on Robotics and Automation*, Alaska.
- Billing, E. A., and Hellström, T., 2008. Behavior Recognition for Segmentation of Demonstrated Tasks, *IEEE SMC International Conference on Distributed Human-Machine Systems*, pp. 228 – 234, Athens, Greece.
- Billing, E. A. and Hellström, T., 2010b. A Formalism for Learning from Demonstration, *Journal of Behavioral Robotics*, Vol. 1, pp. 1–13.
- Brown, A. L., and Cocking, R. R., 2000. *How people Learn: Brain, Mind, Experience, and School: Expanded Edition*, John D. Branford, National Academy Press.
- Byrne R. W., Russon. A. E., 1998. Learning by imitation: a hierarchical approach, *The Journal of Behavioral and Brain Sciences*, Vol. 16.
- Byrne, R. W., 1994. The evolution of intelligence. In *Behaviour and Evolution* (ed. P. J. B. Slater and T. R. Halliday), pp.223-265. Cambridge University Press.
- Coradeschi, S., Saffiotti, A., 2003. An Introduction to Anchoring Problem. *Journal of Robotics and Autonomous Systems*. Vol. 43, pp. 85-96.
- Crestani, F., 1997. Application of Spreading Activation Techniques in Information Retrieval. *Journal of Artificial Intelligence Review*. Vol. 11, pp. 453-482.
- Galindo, C., Saffiotti, A., Coradeschi, S., Buschka, P., 2005. Multi-Hierarchical Semantic Maps for Mobile Robotics. In *Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*. pp. 3492-3497.
- Hartley, R. and Barnden, J., 1997. Trends in Cognitive Science, Vol. 1, No. 5, elsevier Science ltd.
- Huang, Q. Y., Su, J. S., Zeng, Y. Z., Wang, Y. J., 2006. Spreading activation model for connectivity based clustering. *Advances in Information Systems*. Vol. 4243/2006, pp. 398-407.
- Kasper, M., Fricke, G., Steuernagel, K., Puttkamer, E. V., 2001. A behavior-based mobile robot architecture for Learning from Demonstration. *Journal of Robotics and Autonomous Systems*. Vol. 34, pp. 153-164.
- Kopp, S., Graeser, O., 2006. Imitation Learning and Response Facilitation in Embodied Agents. *Proceedings of 6th International Conference on Intelligent Virtual Agents, Marina Del Ray, CA, USA*. pp. 28-41.
- Markou, M., Singh, S., 2003. Novelty Detection: A Review – Part 1: Statistical Approaches. *Journal of Signal Processing*. Vol. 83, Issue 12, pp. 2481-2497.
- Mugnier, M. L., 1995. On Generalization / Specialization for Conceptual Graphs. *Journal of Experimental & Theoretical Artificial Intelligence*. Vol. 7, Issue 3, pp. 325-344.
- Nehaniv C. L., Dautenhahn K., 2000. Of hummingbirds and helicopters: An algebraic framework for interdisciplinary studies of imitation and its applications. *Learning Robots: An Interdisciplinary Approach*. Vol. 24, pp. 136–161, World Scientific Press.
- Nguyen, H., Deyle, T., Reynolds M., and Kemp, C. C., 2009. PPS-tags: Physical, Perceptual and Semantic tags for autonomous mobile manipulation,

Proceedings of the IROS Workshop on Semantic Perception for Mobile Manipulation.

- Nicolescu, M. N., 2003. A framework for learning from demonstration, generalization and practice in human-robot domains. *Doctoral Dissertation*. University of Southern California, Los Angeles, CA, USA.
- Rodriguez, M. A., 2008. Grammar-Based Random Walkers in Semantic Networks, *Knowledge-Based Systems*, Vol. 21, pp. 727-739.
- Rumelhart, D., Norman, D., 1983. Representation in memory. *Technical Report*, Department of Psychology and Institute of Cognitive Science, UCSD La Jolla, USA.
- Vashchenko, N. D., 1977. Concept Formation In a Semantic Network. *Cybernetics and System Analysis*. Vol. 19, No. 2, pp. 277-285.
- Whiten, A., Ham, R., 1992. On the nature and evolution of imitation in the animal kingdom: Reappraisal of a century of research. *Advances in the Study of Behaviour*, Vol. 21, pp. 239-283.



SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

The logo for SCITEPRESS is centered on the page. It features the word "SCITEPRESS" in a large, bold, sans-serif font. Below it, the words "SCIENCE AND TECHNOLOGY PUBLICATIONS" are written in a smaller, all-caps, sans-serif font. The logo is overlaid on a faint, stylized graphic of a graduation cap (mortarboard) with a tassel.