

Measurement of robot similarity to determine the best demonstrator for imitation in a group of heterogeneous robots

Raphael Golombek, Willi Richert, Bernd Kleinjohann, and Philipp Adelt

Abstract Imitation is not only a powerful means to drastically downsize the exploration space when learning behavior. It also helps to align the learning efforts of a robot group towards a common goal. However, one prerequisite in imitation, the decision of which robot to imitate, is often factored out in current research.

In our work we address this question by providing a means to measure the similarity between two robots. Based on this similarity a robot can choose which robot to imitate. The affinity of two robots with respect to imitation is most reasonably measured by calculating their behavioral difference, since the goal of imitation is learning new behavior. This is accomplished by each robot individually constructing an Affordance Network which is a Bayesian network upon its conditional affordance probabilities in the environment. An affordance represents the interaction possibilities an object provides to the robot. These Affordance Networks are then compared with a new metric.

1 Introduction

Imitation is not only a powerful means to drastically downsize the exploration space when learning new behavior [2, 12, 17]. It also helps to align the learn efforts of a robot group towards a common goal. It becomes especially important if a robot is a member of a group of robots who have to accomplish a common task or even several different tasks. The awareness of this has led to the definition of the “big five” questions in imitation, “namely who, when, what, and how to imitate, in addition to the question of what makes a successful imitation” [5]. As recent research has concentrated on the “what” and “how”, the “who” has so far been factored out in current research – either by restricting the imitation process to a one-to-one demonstrator-imitator relationship where the roles of both are clear, or by providing the robots with fixed rules. However, the question of whom to actually imitate plays a role al-

Intelligent Mobile Systems, University of Paderborn / C-LAB, Germany, richert@c-lab.de

ready in early childhood, as shown e.g. by the Psychologist Burnstein [3]: He found out that children imitate more often peers that have similar sex, age or interests.

In our work we address this question by providing a means to measure the similarity between two robots. The affinity of two robots with respect to imitation is most reasonably measured by calculating their behavioral difference, since the goal of imitation is learning new behavior. This is accomplished by each robot individually constructing an Affordance Network which is a Bayesian network upon its conditional affordance probabilities in the environment. Encoded in such a network is the information which capabilities are dependent on which other ones. If, e.g. a robot knows it has the capability A and another robot is capable of A and B and in addition has identified the dependency $A \rightarrow B$, then it would be wise to imitate that robot. Two robots can thus individually learn about each other's Affordance Networks and calculate the difference. When comparing each other the robots can then determine from this difference the degree of behavioral similarity – the more behaviorally similar two entities are the more reasonable it would be for them to mutually imitate beneficial behavior. Depending on the metrics used to calculate the difference an observing robot can decide whether it should a) copy the knowledge of the other robot (if the demonstrator has the possibilities to share it), b) imitate the robot using indirect observation, or c) ignore it.

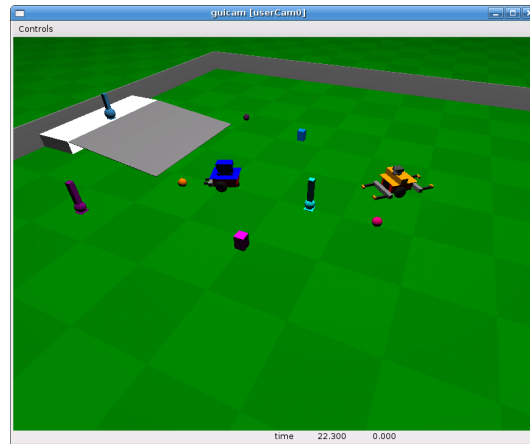


Fig. 1 Simulation environment with two morphologically different robots: Choosing the right robot to imitate helps to avoid useless imitation attempts

2 Related work

The psychologist Gibson observed that our perception of the world is dependent on our interactions with it. For this he introduced the term *affordance* [9, 10] which is a

property an object can have that describes the possible actions that can be performed with it. This depends on the one hand on the object itself, its objective properties like size, weight, surface friction, or shape, and on the other hand on the entity that tries to manipulate that object. But even if the system has all the capabilities to manipulate an object it does not help until it knows how to do that – a common learning problem in developmental robotics [13]: If a system is able to find out which actions make generally sense, it has filtered out the vast amount of useless actions.

Robotics researchers have embraced that concept of affordances as it helps the research to look through the eyes of a robot [7, 15, 16]. Affordances are even used together with Bayesian Networks (BNs) in the field of imitation: Lopes et al. [14], e.g., use BNs to learn object affordances. However, as they are interested in learning the individual affordances they assume the role of the demonstrator to be known. Thus they use BNs to model the affordance. In contrast, we use them to model affordance dependencies to infer behavioral differences. Cesa-Bianchi et al. [4] present an algorithm that chooses the best expert from a set of predefined experts. In their game-theoretic approach they require a static set of always accessible experts – a condition hardly met in realistic robotic domains. As their algorithm is only relying on the experts' performances it does not help in heterogeneous real-world robotics applications, where robots have different morphologies and capabilities. Balch developed a means to measure the overall diversity of a group of robots [1]. In his approach the robots are assumed to be morphologically similar and that their learning algorithms already have converged to a stable behavior.

3 Algorithm

To compare robots based on their behavioural affinity and thus actively control the imitation process one needs data which is related to the robot's behavioural possibilities. Upon the raw data we build a meaningful representation which we then use to compare robots by an adequate metric.

3.1 *Gathering raw data*

First we need data about behavioural possibilities of a robot. This information relies on the robot's hardware and software. Thus the first idea could be to compare these components. However, this approach is infeasible because the information can be exchanged by communication only and a common communication interface cannot be demanded for arbitrary robots. Furthermore, the robot's hardware differs even if it fulfills the same functionality, i.e. one robot could be differentially driven while another one uses omniwheel for locomotion.

As already pointed out, affordances are subjective and tightly coupled to the behavioural capabilities of a robot. Each robot can gather its own affordance data during environmental exploration and can get the affordance data of potential demonstrators by observation. In this paper we assume the set of affordances to be pre-defined and constant. However, this is no restriction to our algorithm as learning affordances at run-time (e.g. via [14]) does not pose a problem to this approach. Let $S = \{s_1, \dots, s_n\}$ be the set of all recognizable affordances and let $L = \{l_1, \dots, l_n\}$ denote all objects in the robot’s environment. We can then define $O = S \times L \times B$ where $B = \{True, False\}$ as the set of all possible results of an affordance test. Furthermore, we define $O_{r_i} \subseteq O$ as the set of results gathered by robot R_i . To simplify the discussion we define the following functions: $bool(o_{ir}) = b_{ir}$, $affordance(o_{ir}) = s_{ir}$, and $object(o_{ir}) = l_{ir}$, whereas $o_{ir} \in O_{r_i}$.

3.2 Affordance Network

The gathered raw data is unstructured, noisy and incomplete. Furthermore, the set of data samples grows rapidly during the robot’s environmental exploration. To structure the data and cope with uncertainty each affordance is interpreted as a random variable X_i and the gathered data for this affordance as a sample set. We then can define the finite set $X = \{X_1, \dots, X_n\}$ of random affordance variables where each variable may take on a value x_i from the domain $\{True, False\}$. Upon O_{r_i} we define $O'_{r_i} = \{(bool(o_{r_1}), \dots, bool(o_{r_m})) \mid \forall o_{r_i}, o_{r_j} : i \neq j \text{ and } object(o_{r_i}) = object(o_{r_j})\}$ and interpret these tuples as samples of the joint distribution of the random variables in X . To get a compact representation of the joint distribution of the variables in X we train a Bayesian Network with the set O'_{r_i} for each robot individually. As the data in O'_{r_i} is directly coupled to the tested affordances a Bayesian Network trained with this set will also encode behavioral information.

Definition 1. Let P be a joint probability distribution of the random variables $X = \{X_1, \dots, X_n\}$ in some set V , and $\mathbb{G} = (V, E)$ be a DAG. We call (\mathbb{G}, P) a Bayesian Network if (\mathbb{G}, P) satisfies the Markov condition. By applying the chain rule of probabilities and properties for conditional independencies, any joint distribution P that satisfies the Markov condition can be decomposed into the product form:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i \mid parents(X_i)).$$

The directed edges of the DAG describe causal relations between the random variables in X . Each node has an attribute which describes the conditional probabilistic distribution of its random variable and the random variables of its parents.

If there is enough expert-knowledge to define the structure of a Bayesian Network, only the parameters i.e. the conditional probabilities have to be learned from data. In our case as we do not use any further domain knowledge we have to learn both, the structure and the parameters from data. Therefore we apply the Structural

EM Algorithm [8] to the affordance data. This is an iterative algorithm based on a standard Expectation Maximization algorithm to optimize parameters, and a structure search to find the current best structure model.

A problem commonly found in structural-learning Bayesian networks is that real causality cannot be derived from raw data [11]. However, this is not a problem here as we do not make inference on the trained networks rather we use them to measure the distance between robots by means of behavioral affinity.

3.3 Metric

After defining a meaningful and well structured representation of a robot we now need to define a metric to measure robot affinity. As BNS are directed acyclic graphs we can apply the Graph Edit Distance Metric (GED) [6] to measure structural distance between two graphs g_1 and g_2 . To describe the GED metric we need the definition of a label representation as defined in [6]:

Definition 2. Let L_E and L_V denote sets of edge and node labels, respectively. A graph $g = (V, E, \alpha, \beta)$ is a 4-Tuple where V is the finite set of vertices, $E \subseteq V \times V$ is the set of edges, $\alpha : V \rightarrow L_V$ is a function assigning labels to the nodes and $\beta : E \rightarrow L_E$. The label representation of g , $p(g)$, is given by $p(g) = (L, C, \lambda)$:

- $L = \{\alpha(x) | x \in V\}$,
- $C = \{(\alpha(x), \alpha(y)) | (x, y) \in E\}$, and
- $\lambda = C \rightarrow L_E$ with $\lambda(\alpha(x), \alpha(y)) = \beta(x, y)$ for all $(x, y) \in E$.

Using the label representation we can then define the Graph Edit Distance metric [6]:

Definition 3. Let g_1, g_2 be two graphs with label representations $p(g_1)$ and $p(g_2)$. Furthermore, let $C_0 = \{(i, j) | (i, j) \in C_1 \cap C_2 \text{ and } \lambda_1(i, j) = \lambda_2(i, j)\}$ and $C'_0 = \{(i, j) | (i, j) \in C_1 \cap C_2 \text{ and } \lambda_1(i, j) \neq \lambda_2(i, j)\}$. Then the graph edit distance $d(g_1, g_2)$ of the two graphs is

$$d_{ged}(g_1, g_2) = |L_1| + |L_2| - 2|L_1 \cap L_2| + |C_1| + |C_2| - 2|C_0| + |C'_0| \quad (1)$$

As we are only interested if an edge between two nodes exists, we define $\beta(x, y) = 1 \forall (x, y) \in E$, and thus omit $|C'_0|$ from our distance metric for affordance networks. As the nodes of the Affordance Networks also need a special treatment due to the inherent conditional probability differences, this leads to the edge comparing part of the final affordance network metric d_{ane} :

$$d_{ane}(g_1, g_2) = |C_1| + |C_2| - 2|C_0| \quad (2)$$

The nodes of the Affordance Network contain conditional probabilities which may differ so that we have to measure distances between these probabilities. For example, consider the node with the label ‘‘Reachable’’ in the networks g_1 and g_2

(Fig. 2). In g_1 it has no parent, thus it is said to be unconditioned. The node with the same label in g_2 is conditioned by the outcome of node with the label “Pushable”. We have to compare $P(\text{Reachable}) = 0.2$ from the node in g_1 with $P(\text{Reachable}|\text{Pushable} = 1) = 0.0$ and $P(\text{Reachable}|\text{Pushable} = 0) = 0.8$ from the node in g_2 . To be able to measure the distance inside the nodes we use the definition

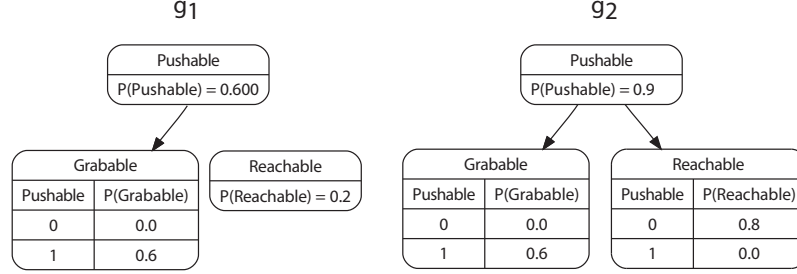


Fig. 2 Two Affordance Networks with different distributions

of independence and conditional independence.

Definition 4. Two events E and F are independent if one of the following holds:

- $P(E|F) = P(E) \wedge P(E) \neq 0, P(F) \neq 0$
- $P(E) = 0 \vee P(F) = 0$

Definition 5. Two events E and F are conditionally independent given an event G if $P(G) \neq 0$ and one of the following holds:

- $P(E|F \cap G) = P(E|G) \wedge P(E|G) \neq 0, P(F|G) \neq 0$
- $P(E|G) = 0 \vee P(F|G) = 0$

Using these definitions we can expand the probability of the node in g_1 : The probability for the “Reachable” affordance is therefore transformed from $P(\text{Reach.}) = 0.2$ to $P(\text{Reach.}|\text{Push.}) = 0.2$ and $P(\text{Reach.}|\neg\text{Push.}) = 0.2$ if the events “Reachable” and “Pushable” are independent. Since there is no edge in g_1 between the nodes with labels “Reachable” and “Pushable” the Markov condition guarantees their independence.

After extending the probability labels we can interpret the probabilities of node v as a point $point(v)$ in n -dimensional space, where n is the number of entries in the nodes probability table. Then we can calculate the distance δ (e.g. Euclidean) between two equally labeled nodes in the different Affordance Networks:

$$d_{ann}(g_1, g_2) = \sum_{\substack{v_1 \in V_1, v_2 \in V_2 \\ \alpha(v_1) = \alpha(v_2)}} \delta(point(v_1), point(v_2)) \quad (3)$$

The final distance function for two Affordance networks d_{an} is then the weighted summation of $d_{ane}(g_1, g_2)$ and $d_{ann}(g_1, g_2)$, where the weights are domain dependent and can be used to control the influence of the structure and the probability

distribution:

$$d_{an}(g_1, g_2) = c_e (|C_1| + |C_2| - 2|C_0|) + \quad (4)$$

$$c_n \left(\sum_{\substack{v_1 \in V_1, v_2 \in V_2 \\ \alpha(v_1) = \alpha(v_2)}} \delta(\text{point}(v_1), \text{point}(v_2)) \right) \quad (5)$$

4 Experimental Results

The presented approach will be demonstrated with two scenarios: In the first one it is shown in detail how the approach leads to the determination of behavioral difference. The second scenario demonstrates how its usage leads to a significant improvement of the imitation process.

4.1 Scenario 1

In this artificial example there are three robots of which one is the imitator (R_i) that has to choose between two demonstrators (R_{d_1}) and (R_{d_2}) to imitate. The properties regarding the gripper and the drive motor are shown in Tab. 1. Fig. 1 shows the two demonstrators to which the environmental objects have different affordances because of their different morphologies: The yellow robot (R_{d_2}) with the barbed gripper, e.g., is able to pull objects and can not lift them, whereas the blue one (R_{d_1}) has a strong gripper but a weak drive, so that it is able to lift some objects, but can not pull them.

Over the course of its lifetime the imitator has recorded the affordances of the two other robots in the scenario and has built an Affordance Network as described in Sec. 3.2 that is depicted in Fig. 3 (page 8). Let us now take a look on how the presented behavioral metric works on those networks to measure the behavioral similarity.

Table 1 Qualitative description of three robots. The imitator is more similar to demonstrator 1 in terms of its gripper and motor capabilities. When imitating another robot in order to learn new behavior it should imitate that robot instead of demonstrator 2.

robot	capabilities			
	gripper			motor
	length	strength	style	strength
demonstrator 1	long	weak	barbed	strong
demonstrator 2	short	strong	normal	weak
imitator	normal	weak	barbed	normal

From the qualitative description it is intuitively clear in this simple example that the imitator has more resemblance to demonstrator 1 and should imitate that robot instead of demonstrator 2 (R_{d_2}). Applying the distance metric d_{an} to the data collected with the three robots we get the results for various weights of the edge and the node distance part as shown in Tab. 2. As can be seen the behavioral distance between the imitator and the first demonstrator is smaller than the distance to the second one.

Table 2 Behavioral similarity calculated using the distance metric d_{an} .

c_e	c_n	$d_{an}(R_i, R_{d_1})$	$d_{an}(R_i, R_{d_2})$
0.1	0.9	2.307	3.614
0.25	0.75	2.423	3.845
0.5	0.5	2.61	4.23
0.75	0.75	2.808	4.615
0.9	0.1	2.923	4.846

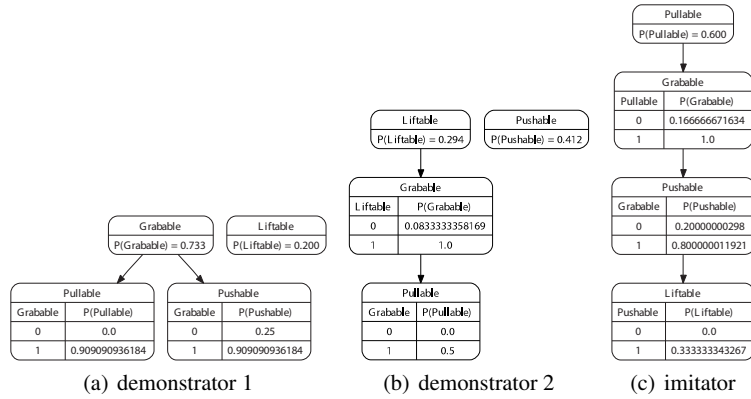


Fig. 3 The final Affordance Networks from the viewpoint of the imitator. These are used by the imitator to calculate the difference in order to determine that robot that has behaviorally the most resemblance to the imitator

4.2 Scenario 2

This scenario was carried out in the PlayerStage/Gazebo simulation environment (Fig. 1). A robot similar to the well-known Pioneer2DX had to choose between three morphologically different demonstrators whom to imitate. All robots differed in the size, strength and shape of their gripper and the strength of their drive unit.

The experiment was carried out as follows: The imitator started without any knowledge and no Affordance Networks. Then it observed a random demonstrator carrying out an action like pushing an object (Fig. 1) and recorded whether it was successful. Afterwards it carried out all known tasks with all known objects in the environment by itself, recorded the success and updated the Affordance Networks for them both. The number of failed behaviors dependent on the number of imitations it has carried out is shown in Fig. 4. As can be seen the imitation process

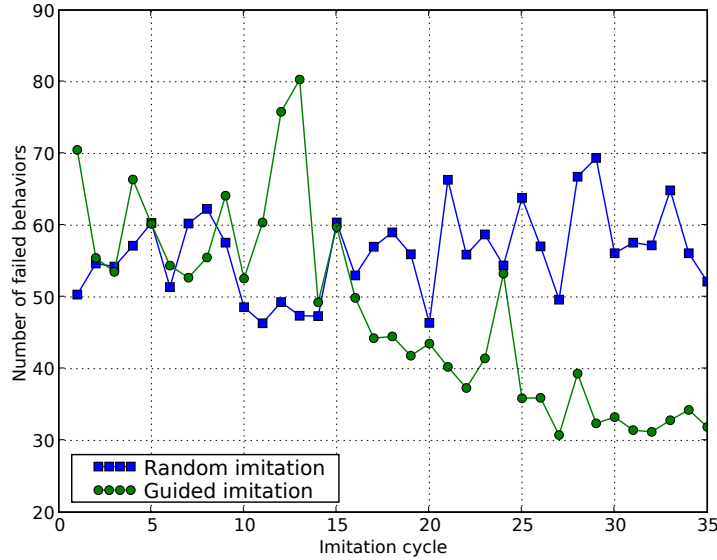


Fig. 4 Increase of the imitation efficiency due to improved imitator selection

gets more efficient with more observations and thus more exact Affordance Networks compared to randomly choosing a demonstrator. The Affordance Networks together with their metric significantly help to improve the overall imitation process.

5 Conclusion

We introduced Affordance Networks, which are Bayesian Networks based on the affordance dependencies of robots and developed a distance metric which calculates the behavioral diversity between two robots based on those Affordance Networks. Once the affordances are observed a robot willing to imitate is able to determine that robot in a heterogeneous robot group that is the most similar to the imitator. This leads to a higher success probability in the imitation process as imitation of robots that are e.g. morphologically different to the imitating robot is avoided.

Although this example is carried out only in realistic simulation environment it should be clear that it would result in a similar outcome in a real world if the data can be collected in sufficient quantity, because the presented approach is able to cope with missing and noisy data. To our knowledge this is the first solution to answer the question whom to imitate in a group of robots. It is not restricted to the robotics domain, but can be applied to all domains where behavior can be observed and imitated.

References

1. Balch, T. *Behavioral Diversity in Learning Robot Teams*. PhD thesis, Georgia Institute of Technology, Dec. 1998.
2. A. Billard and M. J. Mataric. Learning human arm movements by imitation: : Evaluation of a biologically inspired connectionist architecture. *Robotics and Autonomous Systems*, 37(2-3):145–160, 2001.
3. E. Burnstein, E. Stotland, and A. Zander. Similarity to a model and self-evaluation. *Journal of Abnormal and Social Psychology*, 62:257–264, 1961.
4. N. Cesa-Bianchi, Y. Freund, D. Haussler, D. P. Helmbold, R. E. Schapire, and M. K. Warmuth. How to use expert advice. *J. ACM*, 44(3):427–485, 1997.
5. K. Dautenhahn and C. Nehaniv. An agent-based perspective on imitation, 2002.
6. P. J. D. et. al. On graphs with unique node labels. In *Graph Based Representations in Pattern Recognition*, volume 2726, pages 409–437, Heidelberg, DE, 2003. Springer Berlin.
7. P. Fitzpatrick, G. Metta, L. Natale, S. Rao, and G. Sandini. Learning about objects through action-initial steps towards artificial cognition. *Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on*, 3, 2003.
8. N. Friedman. Learning belief networks in the presence of missing values and hidden variables. In *Proc. 14th International Conference on Machine Learning*, pages 125–133. Morgan Kaufmann, 1997.
9. J. Gibson. The theory of affordances. In R. Shaw and J. Brandsford, editors, *Perceiving, Acting, and Knowing: Toward and Ecological Psychology*, pages 62–82. Erlbaum, Hillsdale, NJ, 1977.
10. J. J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton-Mifflin Company, Boston, 1966.
11. D. Heckerman, D. Geiger, and D. M. Chickering. Learning bayesian networks: The combination of knowledge and statistical data. *Mach. Learn.*, 20(3):197–243, 1995.
12. A. Ijspeert, J. Nakanishi, and S. Schaal. Movement imitation with nonlinear dynamical systems in humanoid robots, 2002.
13. M. Kopicki, A. Sloman, J. Wyatt, and R. Dearden. Learning object affordances by imitation. Technical report, The University of Birmingham, 2005.
14. M. Lopes, F. Melo, and L. Montesano. Affordance-based imitation learning in robots. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007.
15. E. Oztop, N. Bradley, and M. Arbib. Infant grasp learning: a computational model. *Experimental Brain Research*, 158(4):480–503, 2004.
16. A. Slocum, D. Downey, and R. Beer. Further Experiments in the Evolution of Minimally Cognitive Behavior: From Perceiving Affordances to Selective Attention. *From Animals to Animats 6: Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior*, 2000.
17. A. Ude, T. Shibata, and C. G. Atkeson. Real-time visual system for interaction with a humanoid robot. *Robotics and Autonomous Systems*, 37(2-3):115–125, 2001.