Learn to Wipe: A Case Study of Structural Bootstrapping from Sensorimotor Experience

Martin Do, Julian Schill, Johannes Ernesti, Tamim Asfour

Abstract— **In this paper, we address the question of generative knowledge construction from sensorimotor experience, which is acquired by exploration. We show how actions and their effects on objects, together with perceptual representations of the objects, are used to build generative models which then can be used in internal simulation to predict the outcome of actions. Specifically, the paper presents an experiential cycle for learning association between object properties (softness and height) and action parameters for the wiping task and building generative models from sensorimotor experience resulting from wiping experiments. Object and action are linked to the observed effect to generate training data for learning a non-parametric continuous model using Support Vector Regression. In subsequent iterations, this model is grounded and used to make predictions on the expected effects for novel objects which can be used to constrain the parameter exploration. The cycle and skills have been implemented on the humanoid platform ARMAR-IIIb. Experiments with set of wiping objects differing in softness and height demonstrate efficient learning and adaptation behavior of action of wiping.**

I. Introduction

The efficiency with which humans perform manipulation tasks in unstructured and dynamic environments is unattained by robotic systems. The key to this remarkable performance lies in the human cognitive capabilities which enable the autonomous acquisition of knowledge by processing complex sensor information and the application of this knowledge to rapidly explore unknown scenes, objects, and actions. Intelligent robots must be able to rapidly create new concepts and react to unanticipated situations in the light of previously acquired knowledge by making generative use of experience utilizing predictive processes. This process is largely driven by internal models based on prior experience (Inside-out). Such robots must also be able to help and learn from others by sharing these generative, experience based theories through teaching and interaction. During development, stimulus driven outside-in and internally driven inside-out processes need to interact with each other at the earliest possible moment to drive the development of cognitive capabilities. The development of such cognitive capabilities has to be embedded in a learning process in order to verify, extend, and revise this knowledge. Hence, in order to make a crucial step towards more autonomy, robots have to be equipped with similar capabilities.

In [1], the concept of *Structural Bootstrapping* has been introduced to address how generative mechanisms which rely on prior knowledge and sensorimotor experience can be implemented in robotic systems and employed to speed up learning. Structural Bootstrapping – an idea taken from child language acquisition research – is a method which provides an explanation of how the language acquisition process in infants is initiated. Hence, in a robotic context, Structural Bootstrapping can be seen as a method of building generative models, leveraging existing experience to predict unexplored action effects and to focus the hypothesis space for learning novel concepts. This developmental approach enables rapid generalization and acquisition of new knowledge about objects, actions and their effects from little additional training data. Entities of the world are represented in form of Object-Action Complexes (OAC) – affordance-based object-action associations that are understood as semantic sensorimotor categories, which are computable (learnable) and storable in a robotic system (see [2]). OACs are related to state-actions transitions and incorporate object as well as action affordances. This allows the specification of actions based object percepts and vice versa enables the grounding of object representations based on the execution and observation of the actions and their effects. Based on the OAC representation, knowledge structures in the form of internal models are generated and intrinsically grounded. The benefit of this knowledge acquisition approach becomes particularly evident on the sensorimotor level where object and action embedded in a situational context are closely intertwined. The experience gained by actively exploring and interacting with the environment, objects and other agents and by observing the effect of actions is characterized by the specific embodiment. Therefore, representations and models emerging from this experience are better adapted to the robot's morphology and more suitable to capture the sensorimotor contingencies than those generated by traditional disembodied methods. The continuous grounding of internal models and representations through exploration provides a suitable basis for prediction and simulation. In this paper, we provide an example for Structural Bootstrapping and demonstrate the validity of the approach on the sensorimotor motor level. Embedded in a learning cycle we show how generative models describing the relation between object properties and action parameters can be learned from experience and how the these models can be used to make predication using internal simulation. More specifically, we show in the context of table wiping task how action parameters can be predicted and adapted based on the object's softness and size.

This work was not supported by any organization

Martin Do, Julian Schill, Johannes Ernesti and Tamim Asfour are with the Faculty of Informatics, Institute for Anthropomatics and Anthropomatics, High Performance Humanoid Technologies, Karlsruhe, Germany {martin.do,julian.schill,asfour}@kit.edu

II. Related works

Several approaches in the literature deal with the problem of exploration-based learning and generative model construction. In the following, an overview on approaches related to the work presented in this paper is given. In [3], an affordance learning framework is introduced which models dependencies between action and object features in the form of a Bayesian Network. Using a set of manipulation actions (grasp, tap, touch) and based on perceived object features the expected effect of an action to be performed could be estimated. In [4], an interactive learning scheme is introduced which allows the identification of object grasp affordances. Grasp primitives represented in the form of a Dynamic Movement Primitive (DMP) are learned from human grasp demonstrations are grounded based the observed effect (grasp successful or not). Towards structural bootstrapping, in [5], an approach is presented for the learning object grasp affordance through exploration. These affordances are represented by grasp densities which are determined based on the visual features (3D edges) of the object to be grasped. The object grasp affordances are grounded and the grasp densities are refined based on exploration and observation of grasping actions performed by the robot. In [6], an approach is introduced which enables a robot to learn a grasping behavior based on initial reflex-like motor primitives. The execution of these primitives at different speeds and the observation of the tactile feedback when touching an object leads to the generation of further behavior primitives. To link the resulting behavior to different intrinsic and extrinsic object properties, the primitives are executed and the observed effects are categorized using the Support Vector Machine (SVR). For the scenario of object-pushing, in [7], a method is proposed which enables a robot to learn goal-directed push-locations on multiple objects. Using a the SVR method a model is learned from explorative pushing which allow the prediction of the effect of certain pushing action considering the current object shape and pose.

III. The Learning Cycle

In order to enable a robotic system to learn and refine sensorimotor knowledge within a developmental process, a learning cycle has to be formalized which incorporates perceptual and motor skills. As suggested in [8], the presented learning cycle consists of four stages. For our work, we define the initial stage to be the exploration stage. Given generalized representations of objects and actions, the robot explores the scene in order to obtain instantiations of both, object and action. The resulting action and object representation A_1 and P_1 form the basis of an experiment which is conducted in the subsequent stage to create data from which concrete experience can be generated. The robot applies the action A_1 and observes its effect E_1 on object, environment, and on the robot itself. In the third stage, based on the data $D = (P_1, A_1, E_1)$ experience is created by grounding and adapting the representations. In the modeling stage, knowledge in the form of internal models f_E and f_A.

In subsequent iterations i with $i > 1$, the grounding is transferred to novel perceived object representation P_i. Using f_A and f_E the parameters for action $\hat{A}_i$ and the expected effect $\hat{E}_i$ for $(P_i, \hat{A}_i)$ can be predicted. $\hat{A}_i$ can be used to constrain and control the exploration of the action parameter space within the repeated experiment and with $\hat{E}_i$ less, however, more relevant additional training data can be created which has to be considered for the re-grounding the representations and revision of the internal models. Hence, this learning cycle allows the continuous acquisition, validation, and refinement of internal knowledge in long term association through exploration and predictive reasoning.

A. *Instantiation of the Learning Cycle for Wiping*

Based on the learning cycle described in Sec. III, a behavior is implemented which enables a robot to efficiently learn wiping movements with different objects. Using skills which have been implemented on our platform, the learning cycle has been instantiated as depicted in Fig. 1. To accelerate the learning process, observations of human wiping demonstrations trigger the bootstrapping process and provide data based on which a coarse representation of the wiping action can be inferred. The wiping action is represented in the generalized form of a periodic DMP (see Sec. IV-A). In the initial iteration, the robot is focused on the adaptation of this representation to environmental circumstances, namely the surface to be wiped. This step corresponds to the grounding of the action representation.

In subsequent iterations, the robot attempts to establish the link between a object, action, and effect. For the object perception, a skill (as described in Sec. IV-B) is applied which enables the robot to deform an object. Based on the extent of the deformations the object's height h and softness s is determined. Thus, a potential wiping object is represented by $(s,h) \in \mathbb{R}^2$. To generate differently scaled wiping movements, the amplitude parameter incorporated in the learned periodic DMP representation can be varied. Especially, regarding the movement of the endeffector directed towards the table, the amplitude has to be scaled according to the specific softness parameter. The search for the optimal amplitude parameter a entails considerable effort since it involves the variation of a, the subsequent parameterization of the wiping primitive, and the reproduction of a wiping action. To assess the effect of a wiping movement the robot visually determines the dirt level (see Sec. IV-C) describing the ratio between the amount of remaining dirt enclosed by an area to be wiped and the entire wiping area size. Hence, the action parameter space is explored for the movement primitive in order to generate a wiping movement with which the wiping success can be maximized. For each stage, a separate experiment is specified. However, the goal for both experiments remains the same: wipe until the dirt level does not change. For a dirt levels $d_i, d_{i-1} \in \mathbb{R}$ determined in iteration i and $i-1$, the goal can be formalized as follows:

$$d_i - d_{i-1} \leq d_\varepsilon \tag{1}$$

where d_ε denote a threshold at which the dirt level change

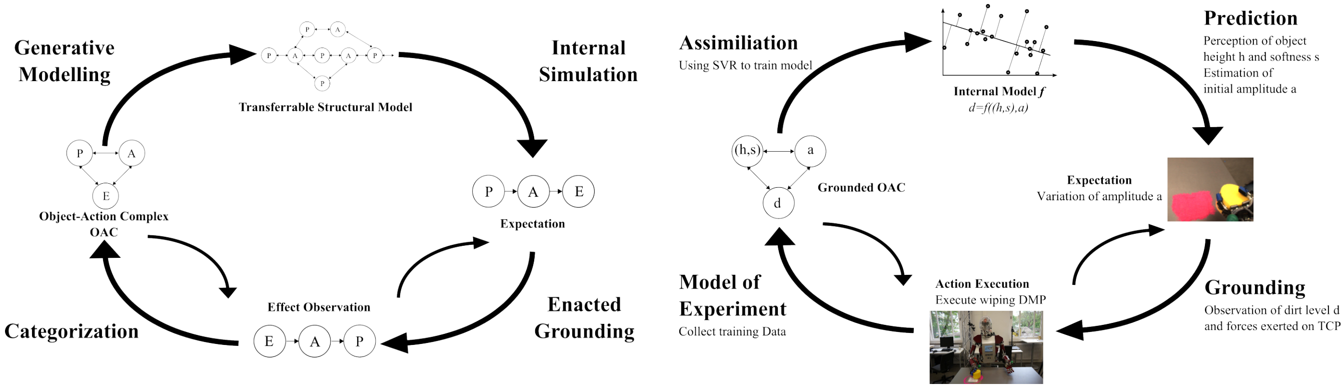

Fig. 1: Left: Abstract learning cycle. Right: Instantiated learning cycle for the learning of wiping.

can be disregarded. To enhance the adaptation of a wiping primitive to novel objects, based on sensorimotor experience gathered in previous iterations, internal knowledge structures are derived. In the form of models, these are used for the prediction of the expected wiping effect for a specific object-action complex. Given a desired effect, these models allow the estimation of the amplitude parameter. Ideally, the action parameter search is conducted in the vicinity of the amplitude estimate.

B. *Surface Adaptation*

The grounding of the wiping DMP corresponds the adaptation of the DMP in order to attain goal-directed wiping movement. In the context of wiping, one prerequisite is constant contact of the object and the surface to be wiped. Therefore, wiping movements can only be adequately evaluated and adapted based on the forces exerted on the robot's end effector. Based on a wiping primitive which encodes a periodic movement pattern p_w in a (x,y)-plane parallel to the surface, we wish to adapt the movement to the shape of the surface. Following the force profile adaptation method introduced in [9], a force-feedback control mechanism is implemented which moves the end effector towards the surface while executing the wiping pattern. In this work, we restrict ourselves to the wiping of flat surfaces. Hence, for a periodic wiping trajectory $p_w(t) = (x_w(t), y_w(t))$ with $T_s < t < T_e$ and T_s, T_e denoting the start and end time of a period, a movement $z_w(t)$ with each discrete time step δt is determined according to following equation:

$$\dot{z_w}(t) = k_f(f_{z_w}(t) - f_0) \tag{2}$$

$$z_w(t) = z_0 + \dot{z}_w(t)\delta t. \tag{3}$$

Here, z_0 stands for the initial height from which the wiping movement is initiated, f_0 denotes the desired force with which the robot should press an object towards the surface, $f_{z_w}(t)$ is the measured force on the end effector, and k_f describes a force gain factor. A further simplification which allows which allows a safer execution of the experiment is to replace f_{z_w} with $f_{z_w} = \sqrt{f_x^2 + f_y^2 + f_z^2}$, since it forces the robot to move upwards when the robots collides with anything from any direction. As a result, the experiment leads to data triplet center of the wiping area $p_0 = (x_0, y_0)$:

$$(P, A, E) = (p_0, (p_w, z_w), d) \tag{4}$$

based on which the action representation is grounded and extended.

C. *Action Parameter Exploration*

To attain an optimal wiping behavior with a specific object, the wiping action has to be parameterized according the object properties. This can be accomplished by specifying the amplitude with which a wiping action is executed. To find a suitable parameterization, the action parameter space is explored within the wiping experiment based on the forces acting on the robot. Starting from an initial estimate a_0, the amplitude is varied according following rules:

$$a(t) = \begin{cases} b^- a(t-1) & , f_{z_w}(t) - f_0 > \rho,\ \dot{z}_w < 0 \\ b^+ a(t-1) & , f_{z_w}(t) - f_0 > \rho,\ \dot{z}_w > 0 \\ b^+ a(t-1) & , f_{z_w}(t) - f_0 < -\rho,\ \dot{z}_w < 0 \\ b^- a(t-1) & , f_{z_w}(t) - f_0 < -\rho,\ \dot{z}_w > 0 \\ a(t-1) & \text{else} \end{cases} \tag{5}$$

where $0 < b^- \leq 1$ and $b^+ = 2 - b^-$ denotes a scalar factors which decreases respectively increase the amplitude according the current movement direction and exerted forces. To accommodate potential noise contaminating the force torque sensor readings, instead of fixating the desired surface pressure on f_0, ρ is introduced into the amplitude update rule to define a range of force values $[f_0 - \rho, f_0 + \rho]$ in which the forces acting on the endeffector are considered to be optimal. $a(t-1)$ represents the amplitude estimate made in the previous time step. For each iteration i, the overall amplitude factor a_i is calculated by $a_i = \frac{1}{T_E - T_S} \sum_{t=T_S}^{T_E} a(t)$. The data which results from the experiments, can be described as follows: for the current object wiping:

$$(P, A, E) = ((s, h), a, d). \tag{6}$$

This data matrix provides the basis for the inference of an internal model.

D. Learning of Internal Models

To generate an internal model representing the relationships between perception, action, and effect, computational methods have to be applied which are suitable to identify structures from non-linear data of arbitrary dimensionality without any prior knowledge. In this work, the Support Vector Regression, a supervised learning technique which is described in [10], is applied to approximate such a model, since it allows to capture complex relationships between the training data points. Furthermore, a sparse model can be obtained by applying the Support Vector method which facilitates the processing of large datasets and enhances the prediction and simulation using the internal model. Based on our experimental data collection $\{(P_n, A_n, E_n)\}_{i=1,...,N}$, for the training of f_E, a dataset D with N input/output pairs is formed as follows::

$$D = \{(x_n, y_n)\}_{n=1,...,N}, \; x_i = (P_i, A_i), \; y_i = (E_i). \tag{7}$$

The internal model is described by $f_E : x \rightarrow y$. Finding a non-linear mapping appropriate function f_E solves the learning problem and leads to desired model enabling the mapping of an arbitrary input pair (P,A) on expected effect $\hat{E}$. Usually, the search for f_E is performed by determining an approximation $\hat{f}_E$ which minimizes the risk functional:

$$R_{emp}\left[\hat{f}_E\right] := \frac{1}{N}\sum_{n=1}^{N} d(\hat{f}_E(x_n), y_n) \tag{8}$$

with $d(f_E(x), y)$ being a distance function to define the relation between the model's output $\hat{f}_E(x)$ and the correct output y. Using the Support Vector method, the non-linear regression problem incorporated in Eq. 8 is transformed into linear problem by introducing a non-linear mapping $\theta : \mathbb{R} \rightarrow \mathbb{R}^{N_h}$ which projects the original dataset D into a feature space of higher dimensionality. Hence, the SVR consists of finding a hyperplane (w,b) which satisfies:

$$g(x,w) = \sum_{j=1}^{N_h} w_j \theta_j(x) + b. \tag{9}$$

To determine a linear model which captures most training samples within an ε-margin, an ε-loss-insensitive functionis defined as follows:

$$L_\varepsilon(g(x,w),y) = \left\{ \begin{array}{ll} 0 & \text{if } |g(x,w)-y| \leq \varepsilon \\ |g(x,w)-y| - \varepsilon & \text{else} \end{array} \right. \tag{10}$$

is introduced into the risk functional. Hence, our goal is to find a function f_E whose distance to any given data point does not exceed ε while being as flat as possible. This optimization problem can be described:

$$\text{minimize } \tau(w) = \frac{1}{2}\|w\|^2 + C\sum(\zeta + \zeta^*) \tag{11}$$

$$\text{subject to } y_i - (g(x_i,w) - b) \leq \varepsilon \tag{12}$$

$$\text{subject to } (g(x_i,w) + b - y_i) \leq \varepsilon \tag{13}$$

where ζ are slack variables which are introduced to the problem in order to relax the constraints and to add a soft margin to the hyperplane and thus to tolerate a small error. C is a constant which controls the trade-off between the flatness of f_E and the tolerated deviations larger than ε. Since θ is unknown according [10] a suitable kernel function such as the Radial Basis Function:

$$k(x,x_i) = \exp(-\gamma\|x - x_i\|) \tag{14}$$

can be used to instead in order to project the data into high-dimensional space. The main parameters controlling the performance of the SVR method are C and the kernel parameter γ.

IV. Implementation

The implementation of the wiping learning behavior is based on skills which already exist on the robot which allow learning and cognition. In the following, the skills and eventual modifications which have been made in order to combine them are briefly described.

A. Wiping Skill

To enable a robot to learn and adapt wiping movements, a skill has been implemented which creates a generalized action representation of a wiping movement. In this work, wiping movements are encoded as periodic DMP using a slight extension of the DMP formulation as suggested in [11] which allows the representation of a periodic motion as well as its corresponding discrete transient movement. In general, a DMP consists of two parts:

$$\begin{cases} \dot{s}(t) = \text{Canonical}(t,s), & (15) \\ \dot{v}(t) = \text{Transform}(t,v) + \text{Perturbation}(s). & (16) \end{cases}$$

The perturbation term in (16) is adapted to a demonstrated trajectory where the transformation system allows the generalization of the learned trajectory to new start and goal conditions. The encoding of both, periodic and transient motion, is accomplished by introducing a two-dimensional canonical system in the DMP formulation: a dimension r to describe distance from the periodic pattern and ϕ denoting the phase of the periodic pattern. This yields the state of the DMP $s(t) := (\phi(t), r(t))$ as the solution (ϕ, r) of the following ordinary differential equation:

$$(17)\begin{cases} \dot{\phi} = \Omega, & (17a) \\ \dot{r} = \eta(\mu^\alpha - r^\alpha)r^\beta. & (17b) \end{cases}$$

Here, $\mu > 0$ denotes the radius of the limit cycle and $\eta, \alpha, \beta > 0$ are constants. The value of $\Omega > 0$ defines the angular velocity of ϕ and has to be chosen according to the period p of the desired trajectory, i.e. $\Omega = \frac{2\pi}{p}$. The value of ϕ is linearly increasing whereas r converges monotonously to μ. Thus, by interpreting (ϕ, r) as polar coordinates the solution of (17) converges towards a circle with radius μ around the origin on the phase plane. To encode a demonstrated wiping action described by $(x_w(t), y_w(t), z_w(t))$, a transformation system in the form of a critically-damped spring system which converges towards a global point attractor g is defined. Therefore, for the encoding of $x_w(t)$ which circulates around

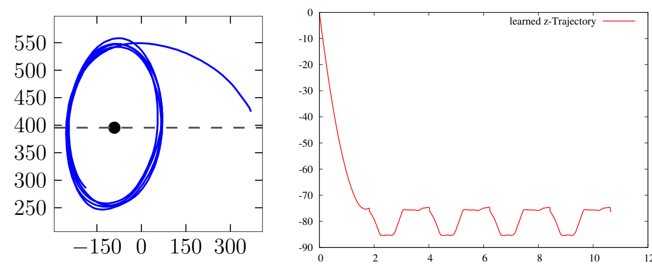

Fig. 2: Left: The wiping pattern p_w extracted from human observation. Right: A generated displacement trajectory z_w.

the attractor $g_x = \frac{\delta t}{T_s - T_e} \sum_{t=T_s}^{T_e} x_w(t)$, the transformation system is specified as follows:

$$\begin{cases} \dot{x}_w = \Omega\Big(\alpha_z\big(\beta_z(g_x - v) - x_w\big) + a \cdot f_x(\phi, r)\Big), \\ \dot{v} = \Omega x_w. \end{cases} \tag{18}$$

The constants $\alpha_z, \beta_z > 0$ are chosen according the ratio $\frac{\alpha_z}{\beta_z} = \frac{4}{1}$ in order to ensure critical damping. By adapting f_x corresponding to the demonstrated trajectory the system oscillates around g_x in a similar manner as featured the demonstration. Here, f_x is defined as

$$f_x(\phi, r) = \frac{\sum_{j=1}^{M} \psi_j(\phi, r)\tilde{w}_{x,j} + \sum_{i=1}^{N} \varphi_i(\phi, r) w_{x,i}}{\sum_{j=1}^{M} \psi_j(\phi, r) + \sum_{i=1}^{N} \varphi_i(\phi, r)}, \tag{19}$$

where $W_x := (w_{x,1}, \ldots, w_{x,N}, \tilde{w}_{x,1}, \ldots, \tilde{w}_{x,M})^T \in \mathbb{R}^{N+M}$ contains the weights which can be adjusted to fit the desired trajectory $x_w(t)$. The basis functions ψ_j encode the transient part of the motion while the periodic part is modeled using ψ_j. The transformation systems $y_w(t), z_w(t)$ are defined analogously. With $f \equiv 0$ the system state v converges to the anchor point $(g_x, g_y, g_z) \in \mathbb{R}^3$. The factor $a > 0$ is changed on-line during the reproduction of the motion to modulate the amplitude.

The learning of a wiping movement is decoupled in two phases: the learning of the wiping pattern from human observation and the adaptation of a wiping movement primitive to the surface to be wiped. In the first phase, motion data representing human wiping demonstrations gets segmented to identify the transient part and the periodic pattern. The weights in (19) are calculated to make the system reproduce the demonstration. Initially, the wiping movement demonstrated in task space is learned in the (x, y)-plane disregarding the surface contact which yields a wiping DMP with two transformation systems.

In the second phase, an additional transformation system is learned which encodes the movement $z_w(t)$ needed for the adaptation to the surface. To obtain $z_w(t)$, the wiping DMP is repeatedly reproduced until the force torque measurements during the execution of $z_w(t)$ meet predefined constraints which guarantee that the endeffector applies a specific pressure on the surface to be wiped. In Fig. 2, the trajectory $(x_w(t), y_w(t))$ which features the periodic wiping pattern as well as trajectory of the $z_w(t)$ are depicted.

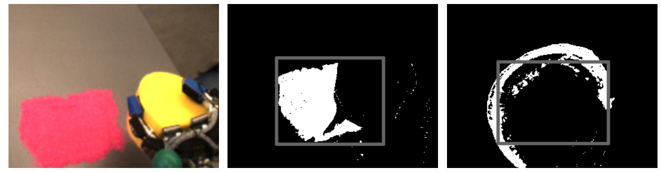

Fig. 3: Left: Robot view on the scene. Center: Segmented view of the scene in the beginning of the wiping execution. Right: Segmented view on a ”clean” table.

B. Softness Skill

To check the deformability and softness of an object the robot uses his ability to control the grasping force of the pneumatic actuated hand with a model based force position control [12]. When the object is in the hand and grasped between the fingertips with a low grasping force, the distance between the fingertips of the index finger, middle finger and thumb is measured using the joint encoders and the forward kinematics. Then the grasping force is increased which results in a deformation of the object. After the fingers have stopped moving, the distance between the fingertips is measured again and the difference of the distances is used as a measure for the softness of the object.

C. Dirt Level Skill

As mentioned before, the effect of a wiping action is described by the dirt level d within area O to be wiped. For the sake of simplicity, it is assumed that dirt features a specific color. Therefore, to determine the size and position of O, using the stereo camera setup the robot explores the table and performs a color segmentation in order to localize the largest blob. A bounding box B_i around that blob provides the image coordinates of O. Transformed into the world coordinate system, one obtains B_w which provide the global coordinates of O. In order to determine the current dirt level at any time t during the execution of the wiping experiment, B_w is transformed back onto image coordinates B_i^t. Hence, based on B_i^t d one can calculate according following equation:

$$d = \sum_{i=y_{min}}^{y_{max}} \sum_{j=x_{min}}^{x_{max}} \frac{k(i,j)}{(x_{max} - x_{min})(y_{max} - y_{min})}. \tag{20}$$

Since the hand might occlude a considerable area of the surface, a reliable assessment of the dirt level cannot be performed at guaranteed at any time during the execution of a wiping movement. Hence, to control the experiment, the current dirt level at t_c is set to $d(t_c) := d_{max,i}$ which is defined as follows:

$$d_{max,i} = \max\{d_i(t)\}_{T_{s_i} < t < t_c} \tag{21}$$

with i denoting the index of the current period. The experiment is finished when following conditions are fulfilled as described by

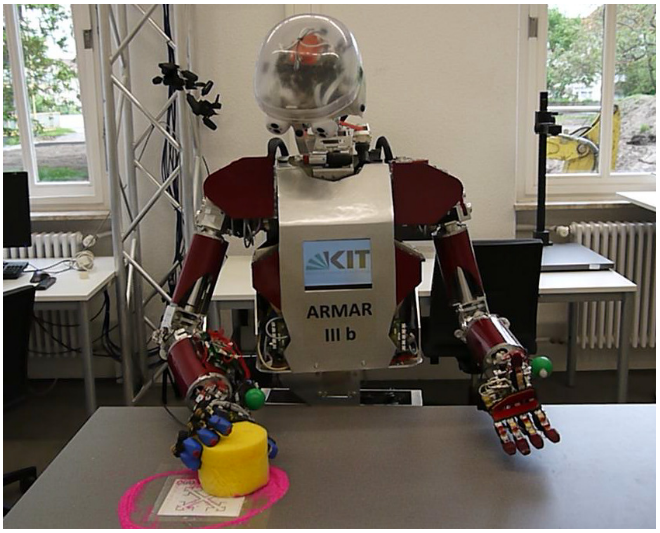

Fig. 4: The humanoid platform ARMAR-IIIb wiping the table with a sponge.

V. Experiments

As depicted in Fig. 4, the implemented learning behavior has been evaluated on our humanoid platform ARMAR-IIIb (see in[13]). The learning of the wiping primitive in the initial iteration is described in. In subsequent iterations, to facilitate the environmental perception, the color of the dirt (pink sand) has been specified. Based on this information, the robot initializes each learning iteration by localizing the dirty area O. The corresponding bounding box B_i is used to specify the target configuration of the DMP. In the following step, the robot determines the object softness and height by grasping the object of interest at the bottom and top side of the object. The object exploration process is assisted by human operator since for wiping the object has to be reoriented in the robot's hand, so that the object is grasped from the side enabling the bottom to touch the table. Given the internal models, predictions are made for the amplitude and the expected effect. Subsequently, the robot performs a wiping movement with a and compares the observed effect with the expected effect. If the observations does not coincide with the expectation, a parameter exploration procedures as described in Sec amplitude is initiated in order to create further data for the grounding of the internal models. For now, the grounding of an internal model is done by updating the data set and retraining the entire model. We are aware that the learning cycle has to incorporate an incremental learning algorithm in order to be effective for the longer term and with an increasing amount of data.

Therefore, in this section, results of preliminary experiments are presented showing the effectiveness of the experience learning cycle for the implementation of a cognitive learning behavior for robots, in particular, in the context of wiping. The wiping experiments have been conducted on set of twelve objects which includes instances designed for wiping (sponges, towel, toilet paper) and other household items (box, bottle, ball, can) that are less suitable. We restrict ourselves on objects whose height and weight are within are within a predefined range in order to prevent damage to the robot. Based on experimental data originating from wiping experiments with this object set, internal models f_E and f_A are generated using the SVR method. In this work, we used the LIBSVM library introduced in [14] for the training. The relevant parameters for the training of f_A have be determined to be $C = 50$ and $\gamma = 0.5$ whereas f_E has been trained with $C = 10$ and $\gamma = 0.33$. The data and predictions of the amplitudes and the expected dirt levels are listed in Table I. It is interesting to see that for soft objects the amplitude could be reliably re-estimated. The main reason for the variation of the amplitudes for harder objects lies in the increased sensitivity towards forces exerted on the object respectively the end-effector. A slight difference of the object pose in hand can produce very different results. Regarding the prediction of the expected dirt level, good estimations could be made for cubic objects. For spherical and cylindric objects, less useful predictions have been inferred.

Given a percept of a specific object, the corresponding amplitude estimate can be used to considerably reduce the adaptation effort of a wiping movement. The plots depicted in Fig. 5 indicate that with increasing knowledge leading to more accurate estimations of the action parameter the execution of an action converges faster towards the desired behavior. With regard to the forces exerted on the end-effector, a force trajectory is desired which oscillates around the predefined force threshold of $f_0 = 25$ whereas regarding the dirt level we wish to minimize the dirt level as fast as possible. The learning phase denotes the initial phase where the movement primitive is adapted to the environment. In the adaptation phase, based on a default value of $a_0 = 1$ the amplitude is varied in order to attain the desired effect. In the execution phase, the task is performed using the estimated amplitude parameter and without any adaptation.

VI. Conclusion

An approach for the implementation of a cognitive learning behavior enabling robots to create individual knowledge structures based on experience gained through physical exploration, interaction, and observation has been proposed. The behavior manifests in the form a learning cycle which

Object	h	s	$\hat{a}$	a	$\hat{d}$	d
sponge (s)	79	0.0343	1.0	1.0	0.162	0.117
sponge (m)	91	0.0384	0.957	0.948	0.129	0.132
sponge (l)	102	0.0358	1.13	1.139	0.139	0.09
styrofoam cube (s)	87	0.00474	0.701	0.696	0.270	0.258
styrofoam cube (l)	91	0.00774	1.0	1.0	0.13	0.177
rolled towel	89	0.0213	1.215	1.057	0.229	0.142
styrofoam ball	100	0.00639	1.497	1.496	0.384	0.422
cardboard box	91	0.0171	1.072	1.072	0.240	0.178
plastic bottle	87	0.0263	1.453	1.453	0.310	0.568
metal can	86	0.00843	1.04	1.366	0.352	0.529
toilet paper	101	0.0232	0.887	0.887	0.162	0.128
foam	91	0.041	0.999	1.0	0.13	0.177

TABLE I: Object properties and the corresponding action and effect parameter. h denotes the object height in mm and s the softness of an object. $\hat{a}$ and a represent the estimated and the actual amplitude of an adapted wiping movement. $\hat{d}$ and d stand for the expected and actual dirt level which indicates the effect of wiping.

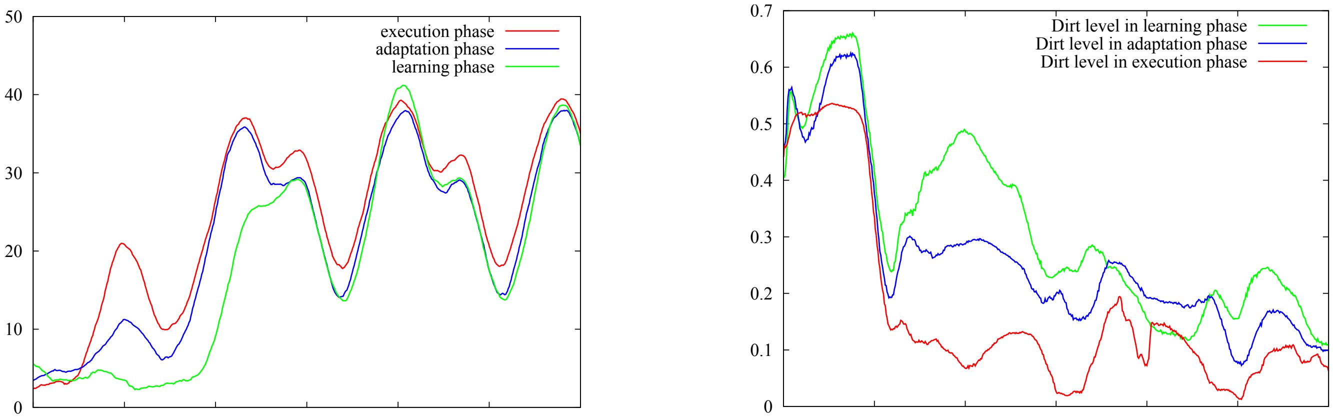

Fig. 5: Left: Trajectories of forces exerted on the end effector in various phases of the wiping learning cycle. Right: Dirt level evolution in various phases of the wiping learning cycle.

incorporates perceptual and motor skills in order to continuously acquire data based on which internal models are generated and grounded. For the scenario of table wiping, we have showed that with these internal models wiping primitives can be efficiently learned and adapted to different task and object-specific constraints.

However, we have also experienced cases in which the estimation of action parameters and the prediction of the expected effect failed. This is mainly due to the simple object representation which merely relies on the object softness and height. As indicated in our results, for deformable objects, these features might suffice in order to determine the affordance in the context of wiping. By redefining the experiment, the application of the for the learning and the adaptation of other actions is limited to actions that are mainly controlled by the amplitude and for which the object softness has a tremendous effect on the outcome of an action such as kicking or throwing. Therefore, a more universal implementation of the structural bootstrapping approach is attained by extending the action parameter space and by incorporating an enriched visuo-haptic object representation which considers further object properties have to be considered such as geometry and weight. Therefore, in the future we will focus on the integration of an enriched object representation which allows the estimation of further action parameters such as different hand orientations or wiping patterns. Furthermore, we will conduct extensive experiments with numerous objects with the goal of enabling a robot to extend the knowledge structures.

Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme under grant agreement no. 270273 (Xperience).

References

[1] "Xperience Project," Website, available online at http://www.xperience.org.

[2] N. Krüger, C. Geib, J. Piater, R. Petrick, M. Steedman, F. Wörgötter, A. Ude, T. Asfour, D. Kraft, D. Omrčen, A. Agostini, and R. Dillmann, "Object-action complexes: Grounded abstractions of sensorimotor processes," *Robotics and Autonomous Systems*, vol. 59, pp. 740–757, 2011.

[3] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, "Learning object affordances: From sensory motor coordination to imitation," *Transactions on Robotics*, vol. 24, no. 1, pp. 15–26, February 2007.

[4] O. Kroemer, E. Ugur, E. Oztop, and J. Peters, "A kernel-based approach to direct action perception," in *IEEE International Conference on Robotics and Automation (ICRA)*, St. Paul, USA, May 2012, pp. 2605–2610.

[5] R. Detry, D. Kraft, O. Kroemer, L. Bodenhagen, J. Peters, N. Krüger, and J. Piater, "Affordance prediction via learned object attributes," *Journal of Behavioral Robotics*, vol. 2, no. 1, pp. 1–17, March 2011.

[6] E. Ugur, E. Sahin, and E. Oztop, "Self-discovery of motor primitives and learning grasp affordances," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, October 2012, pp. 3260–3267.

[7] T. Hermans, F. Li, J. M. Rehg, and A. Bobick, "Learning stable pushing locations," in *IEEE International Conference on Developmental Learning and Epigenetic Robotics (ICDL-EPIROB)*, Osaka, Japan, August 2013.

[8] D. Kolb, *Experiential learning: experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice Hall, 1984.

[9] A. Gams, M. Do, A. Ude, T. Asfour, and R. Dillmann, "On-Line Periodic Movement and Force-Profile Learning for Adaptation to New Surfaces," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Nashville, USA, December 2010.

[10] A. Smola, "Support Vector Regression," *International Journal of Robotics Research*, vol. 30, pp. 1229–1249, September 2011.

[11] J. Ernesti, L. Righetti, M. Do, T. Asfour, and S. Schaal, "Encoding of periodic and their transient motions by a single dynamic movement primitive," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Osaka, Japan, December 2012, pp. 57–64.

[12] A. Bierbaum, J. Schill, T. Asfour, and R. Dillmann, "Force Position Control for a Pneumatic Anthropomorphic Hand," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Paris, France, 2009.

[13] T. Asfour, K. Regenstein, P. Azad, J. Schröder, N. Vahrenkamp, and R. Dillmann, "ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control," in *IEEE/RAS International Conference on Humanoid Robots (Humanoids)*, Genova, Italy, December 2006, pp. 169–175.

[14] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.