

Staged Development of Robot Skills: Behavior Formation, Affordance Learning and Imitation with Motionese

Emre Ugur, *Member, IEEE*, Yuki Nagai, *Member, IEEE*, Erol Sahin, *Member, IEEE*,
and Erhan Oztop, *Member, IEEE*

Abstract—Inspired by infant development, we propose a three staged developmental framework for an anthropomorphic robot manipulator. In the first stage, the robot is initialized with a basic reach-and-enclose-on-contact movement capability, and discovers a set of behavior primitives by exploring its movement parameter space. In the next stage, the robot exercises the discovered behaviors on different objects, and learns the caused effects; effectively building a library of affordances and associated predictors. Finally, in the third stage, the learned structures and predictors are used to bootstrap complex imitation and action learning with the help of a cooperative tutor.

The main contribution of this paper is the realization of an integrated developmental system where the structures emerging from the sensorimotor experience of an interacting real robot are used as the sole building blocks of the subsequent stages that generate increasingly more complex cognitive capabilities. The proposed framework includes a number of common features with infant sensorimotor development. Furthermore, the findings obtained from the self-exploration and motionese guided human-robot interaction experiments allow us to reason about the underlying mechanisms of simple-to-complex sensorimotor skill progression in human infants.

Index Terms—affordance, developmental robotics, imitation, motionese, goal emulation, sensorimotor learning

I. INTRODUCTION

In the last two decades robotic studies inspired by the developmental psychology have gained wide popularity [1, 2]. The major driving force behind this is the notion that adopting a developmental pathway similar to those of infants should pave the road for intelligent and human-like behaving robots. In general, developmental robotics aims to advance science at two fronts [1, 3, 4, 5]: Creating intelligent behavior via mimicking infant development, and understanding human development through modeling and testing the hypotheses from Psychology and Neuroscience. As such, embodiment, sensorimotor learning, exploration and interaction with the environment are often at the core of developmental robotics.

In this paper, we propose a staged developmental skill acquisition framework that transforms an arm-hand robotic system with a limited sensorimotor capacity into a robot

that can perform a set of goal-directed actions, use learned affordances to understand its environment, make predictions about the consequences of its actions, and finally engage in action imitation through possibly multi-step planning.

Discovering behavior primitives and learning the affordances provided by the environment, and learned behavior primitives are central for the proposed framework. We suggest that this early skill development can be achieved without supervision by having an embodied agent interact with its environment. During this interaction the agent organizes its continuous sensorimotor space into discrete sensorimotor representations or schemas (motor primitives, perception mechanisms and predictors) [6, 7] with continuous parameters where necessary. The result of this organization then can be used for simple planning and imitation. In this paper, we report the realized developmental progression of our robot until this point; however, it is possible to take this a step further towards high-level intelligence by associating the representations formed with symbols and operators with which inference, reasoning and communication mechanisms can be built.

In the proposed developmental progression, we envision three major stages as depicted in Fig. 1. In the first stage the robot is assumed to have only two basic movement mechanisms: a basic finger enclose behavior akin to infants' palmar grasp reflex, and one basic arm action, that generates simple arm movements to transport the hand to a salient area in the visual field of the robot, i.e. to the vicinity of an object. These design choices are well supported by the infant literature. In the newborn infant, a basic neuro-muscular infrastructure for reaching and grasping is present: when an object is placed in the palm of a newborn, the tactile stimulation triggers a finger flexion reaction. Similarly, reaching movements aimed towards objects in the center of the visual field are often expressed in newborns' [8, p.235].

Stage I includes the exploration of the parameter space of the arm movement shaped by hand-object contact information. The contact information, if any, sensed during the approach and possibly after the finger enclosure, is used to cluster the executed movements, yielding a set of behavior primitives such as 'push' and 'grasp'. Stage II is concerned with developing visual perception and prediction ability by using the motor primitives developed in Stage I. The robot probes the environment to learn about relations between objects, behavior parameters, and the effects created. The resultant knowledge, i.e. learned affordances and effect predictors, are used by the

E. Ugur is with University of Innsbruck, Institute of Computer Science, Innsbruck, Austria. E. Sahin is with KOVAN research laboratory, Computer Engineering Dept., Middle East Technical University, Ankara, Turkey. Y. Nagai is with Emergent Robotics Lab., Graduate School of Engineering, Osaka University, Japan. E. Oztop is with Ozyegin University, Computer Science Department, Istanbul, Turkey and Advanced Telecommunications Research Institute (ATR), Dynamic Brain Imaging Dept, Japan.

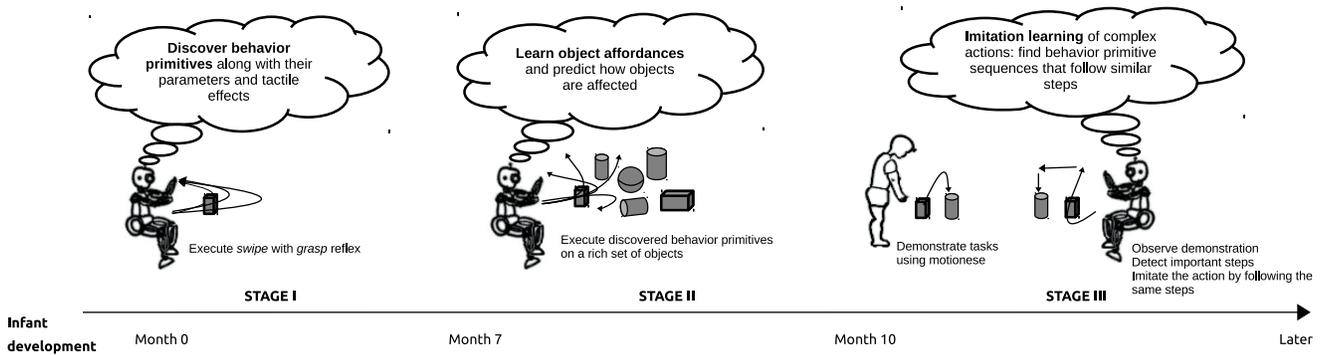


Fig. 1: The staged development of the robot and the caregivers role are summarized.

Please watch the video at <http://emreugur.net/tamd2014/> to see how this development is realized in the real robot.

robot to make plans to achieve a given desired goal state (e.g. a goal can be ‘make the ball disappear’ encoded as a feature vector). In our previous work, we showed that the robot can successfully imitate expert demonstrations by sequential emulating the displayed subgoals [9]. In the expert tutor case, the tutor knew the behavior repertoire, affordance detection and effect prediction capabilities, and imitation mechanisms of the robot. However, a more natural developmental setting should encourage even naïve tutors to adapt their teaching strategy and learn to display motionese for the successful imitation of the robot. With the current work, we are extending those results with new experiments where we employed naïve tutors and analyzed their teaching performance. Regarding the evolution of teaching strategies, although one can expect that the subjects would modify their task demonstrations over several trials, “how” they would modify certain features of the demonstration, such as shape, speed and amplitude of the trajectory is not trivial to predict. To see how naïve subjects engage in motionese in our robotic setting, we employed three naïve tutors who were not informed about the perception and imitation mechanisms of the robot

We suggest that a robot can learn new skills via imitation learning¹ by extracting the important steps from the observed movement trajectory, and then encoding them as subgoals that it can fulfill. In imitation tasks involving objects, replication of the important steps of the observed action is not trivial since objects may have different affordances for humans and robots. As such, a demonstrated action may not correspond to any behavior primitive developed so far, and so its end result cannot be achieved using the goal satisfaction ability provided by Stage II. It is likely that infants also have similar difficulties in mapping observed actions to their own repertoire, and so are not good (complex) imitators during the first natal year. To speed up skill acquisition via imitation, parents and caregivers are known to make modifications in infant-directed actions, i.e. use ‘motionese’ [10], to help infants ‘parse’ the demonstrated actions easier. With the aim of creating this dynamic interaction between the learner and the demonstrator, we implicate Stage III in segmenting, with the help of the tutor,

¹Generally speaking, imitation refers to finding the behavior sequence that enables the robot to follow a similar trajectory as the demonstration. Goal emulation on the other hand refers to computing a behavior sequence to achieve the goal regardless of the followed trajectory.

the demonstrated behavior into chunks that can be replicated by Stage II movement generation mechanism (i.e. goal emulation). This effectively means that the robot can learn a new skill via imitation provided that it can interpret a demonstrated skill as a sequence of subgoals that each can be satisfied in the detected order. If the demonstrator is informed about the kind of motionese she is supposed to display, she can deliberately insert the motionese cues into her trajectory in the ‘right’ moments, enabling the detection and sequential emulation of the subgoals. In our previous work, we showed that the robot can successfully imitate expert demonstrations by emulating the displayed subgoals [11]. In the expert tutor case, the tutor knew the behavior repertoire, affordance detection, effect prediction, and imitation mechanisms of the robot. However, a more natural developmental setting should encourage even naïve tutors to adapt their teaching strategy, and learn to display motionese for the successful imitation of the robot. With the current work, we are extending those results with new experiments where we employed naïve tutors and analyzed their teaching performance. Regarding the evolution of teaching strategies, although one can expect that the subjects would modify their task demonstrations over several trials, “how” they would modify certain features of the demonstration, such as shape, speed and amplitude of the trajectory, is not trivial to predict. To see what is really happening in our setting, for the analysis and evaluation of Stage III, we employed three naïve tutors who were not informed about the perception and imitation mechanisms of the robot, and analyzed their teaching performance and strategies in the current study.

In the next section, we first discuss human infant development with particular attention to the stages that inspired the work presented in this paper, and then summarize the related robotics literature. In Section III, we describe the proposed three stage developmental framework, with details of behavior and affordance representation, learning methods, prediction mechanisms, and imitation. In Section IV, the results of the realized developmental progression on a real robot are provided with the details of discovered behavior primitives, learned affordances and goal-emulation performance. The limitations of the current system, and future research directions are indicated in the Discussion and Future Work Sections. Finally, the Conclusion Section summarizes the results obtained, and

underlines our contribution.

II. RELATED STUDIES

A. Infant development

1) *Early motor development:* The sensorimotor development in humans already starts in the womb [12] and progressively shapes infant behavior after birth into the childhood [13, 14]. Newborns have several innate reflexes such as pupil reflex to light, sucking reflex or palmar-grasp reflex that help the development of motor and cognitive skills. Palmar reflex in particular is “integrated into later intentional grasping” [14, p. 7] after repeated activation of the reflex and execution of grasp action. This reflex is transient, and disappears by 6 months of age [15, p. 199]. By 4 months of age, infants learn to perceive the reachability boundaries [15, p. 199] and they can successfully reach for objects [16, p. 41]. By 5 months of age, infants slow down their hand speeds when grasping objects, i.e. they have learned to adjust hand reach speed by this age [15, p. 100]. It takes 9 months for infants to reach for objects with correct hand-orientation and adjust their grip size based on objects’ size before contact [15]. Hand orientation and grip size appear to develop later than hand-speed parameter since “babies younger than 9-months lack a fully-developed map between visually perceived orientations and corresponding hand orientations” [15, p. 200]. Between 7-9 months, babies explore the environment and objects using various behaviors including grasp, drop, and hit [1]. This indicates that, by this time, the infant has already transformed their initial seemingly uncontrolled ‘move hand’ actions into a set of behavior primitives from its most basic movement primitive, ‘move arm’. Between 7-9 months, they learn the causality relations and object dynamics in response to their actions [1]. It is plausible to think that while interacting with the environment, babies monitor these consequences of their actions and relate the consequences to the visual properties of the objects they interact with. In other words, they learn object affordances, the action possibilities offered by their environment [17], in this stage. Finally, by 10-12 months, they can imitate actions and generate multi-step plans to accomplish goals such as reaching a distant toy resting on the towel by pulling the towel first [18].

Infants between 7-10 months have already acquired a set of behaviors that are qualitatively different and that can be used for different purposes such as grasping, dropping, reaching, shaking, etc. These actions can be considered as behavior primitives that are utilized to develop more advanced skills through practice. There is evidence that complex behaviors are represented in a modular fashion by the central nervous system. For example, the ‘transport’ and ‘grasp’ components of grasping action appear to be controlled by different regions of the human brain [15, p. 217]. Furthermore, there is a developmental order in maturation of these areas. Thus, it is plausible that the infant starts from a small number of reflex-like behaviors, and then progressively discovers and distinguishes new behaviors through the use of existing ones. Such developmental progression must be complemented in infant’s perceptual system. First, a crude perception system

may suffice to discover the basic behaviors, but a more advanced perception is needed to differentiate more complex behaviors and to discover more abstract concepts.

2) *Affordances, means-end behaviors, prediction:* At 7-10 months, the infant starts observing the effects of her hitting, grasping and dropping actions on objects more often, and can learn the dynamics of the objects [1]. The infant in this stage has already acquired a number of manipulation behaviors and is able to detect different properties of objects such as shape, position, and size. Using her motor skills, she interacts with the environment accumulating knowledge about the relationships between objects, actions and the effects that she perceives.

According to Elsner and Hommel[19], infants learn to use anticipation for goal-directed actions in two phases. They execute random actions in the environment, self-monitor the changes, and learn the action-effect associations in a bi-directional way. Later, they start to control their actions by predicting the effects they can create. This process effectively corresponds to the learning of the affordances[17] provided by the environment. The learning in this stage is largely performed in a goal-free fashion through self-exploration and self-observation [20, 21, 22, 23]. After approximately 9 months of age, the infant starts using the learned object-action-effect relations in a goal-directed way, anticipating a desirable change in the environment and behaving accordingly [24, 25, 26]. Behavior in this developmental period involved recalling action-effect mappings and making simple plans that may involve multiple steps [18]. Goal emulation, a form of imitation characterized by the replication of the observed end effect [23], starts after this period, and infants become skilled at emulating unseen movements after 12 months of age [27].

3) *Goal emulation, imitation and motionese:* Infants’ means of imitation changes over time. While younger infants are more inclined in achieving the goal of a demonstrated action, older infants tend to exactly imitate (and in later stages over-imitate) the observed target action sequence even if those actions are not physically related to the goal [28]. Imitating an action sequence is difficult for young infants as they need to map the observed actions to their own sensorimotor repertoire [29] (also see [30] for neurophysiological evidence on how infants’ own action repertoire affects the understanding of observed actions of others.). Thus, a big challenge for successful imitation performance is to map the observed actions onto the observer’s motor repertoire.

In order to deal with this challenge, parents support infants by making modifications in infant-directed actions, i.e. use “motionese” [10, 31]. Motionese is characterized by higher range and simplicity of motion, more pauses between motion segments, higher repetitiveness of demonstration, and more frequent social signals to an infant [10, 32]. Fine-grained analysis using a computational attention model further reveals the role of motionese in action learning [33]. Longer pauses before and after the action demonstration underline the initial and final states of the action (i.e. the goal of the action) whereas shorter but more frequent pauses between movements highlight the subgoals of the action [34]. Of particular interest is that such modifications are elicited by the responses of an action learner [35]. Not only the age of a learner but also

her ability to recognize the demonstrated action (e.g. visual attention) influences the task demonstration.

In this section we focused on imitation that involves manipulation of objects, and discussed that this form of imitation is observed in later stages of infancy. On the other hand, object-free imitation, i.e. imitation involving only body parts, has been reported to be present in newborns [36]. Whether this early form of object-free imitation shares common mechanisms with the latter object based imitation is still an open question in biological systems [37], which we do not consider in this paper.

B. Developmental robotics literature

Recently, the role of proprioceptive and tactile perception in early motor development has been studied in different hand-arm robot systems. Although the crucial role of proprioceptive and tactile perception in human for grasping and manipulation has been long known, it is only recently that these have been incorporated in robotic studies. Oztop et al. [38] emphasized the importance of tactile feedback and its precedence over vision during grasp development, and developed a computational model that parallels human grasp learning. A range of robotics studies also have shown that tactile alone and in combination with vision enables dexterous manipulation in hardware (e.g. [39]). To mimic infant learning, Savastano and Nolfi [40] proposed a neurobotic model where a simulated humanoid robot learns reach and grasp skills incrementally. Similar to our work, they use simple reactive behavior models to explore a given object using touch. In the experiments, they systematically manipulate maturation constraints leading to various development strategies, and observe similar characteristics with infant development for example in motor babbling strategies. Although these studies provide more psychologically plausible models where strategies obtained during development match those of infant strategies, they span a relatively short period of development, and all were realized in simulation. On the other hand, in Saegusa et al. [41], a real humanoid robot learns coordination of head, arm and finger joints through motor babbling for fixation, reach and grasp actions based on sensory feedback. Unlike our developmental system where behavior primitives are self-discovered, they used pre-defined target configurations for these actions, and their focus was in learning the association between actuator commands and visual, proprioceptive and somatosensory effects. Tani and Ito [42] studied behavior primitive formation within neurodynamical systems framework, where diverse set of non-trivial emergent behavior patterns can be generated by modulating so-called parametric bias (PB) parameters of a recurrent neural network. This framework was realized in a low-dimensional robot arm-hand system with visual and proprioceptive sensors obtaining a number of end-point and oscillatory behavior patterns. Built upon [42], Nishide et al. [43] models robot and object motion experience obtained from push actions on cylindrical objects. The system self-organizes experience based on motions of the objects, and provides object-posture dependent push regions that would generate reliably predictable rolling motions. Our system critically explores hand-closure, which generates richer interactions and leads to more complex sensorimotor skills.

Learning affordances in the form of object-action-effect relations has been widely studied in robotics in recent years [44, 45, 46, 47, 48], and has been reviewed in [9]. Overall it can be stated that the affordances learning framework that we have proposed and develop further in this paper, provides the following three characteristics that do not exist simultaneously in any other previous work: (1) multi-step planning, (2) categorization of the perceptual space based on actions and their effects with continuous parameters, and (3) generalization of the knowledge obtained through exploration. A number of studies address different combinations of these three capabilities, however seamless and complete integration of these characteristics is still missing as discussed in detail in [9]. More recently, affordance exploration was combined with intrinsic motivation guided learning in [49], was linked to language learning in [50, 51], and was integrated into human activity learning in [52]. While Ivaldi et al. [49] do not particularly focus on affordance learning, in their experiments, the iCub humanoid robot learns object properties by actively choosing among objects to explore, actions to execute and caregivers to interact. This socially guided intrinsic motivation framework [53] that combines robot's manipulatory actions with social guidance significantly increases object recognition performance, and can be directly used to increase speed of affordance learning. In Koppula et al. [52], object affordances are learned along with human activities in an integrated framework. This system is modelled with a Markov Random Field where the nodes encode the subactivities and affordances; and the edges correspond to the learned relations between these components. Based on RGB-D video input, the system was realized using PR2 robot where not only initial/final perceptual state but also temporal evolution is studied.

The idea of using previously developed capability for affordance prediction in imitation learning is not new. Lopes et al. [54] also used affordances that are modelled in Bayesian Networks to interpret demonstration and to imitate with the robot's own behaviors. While they were able to recognize one-step actions and make one-step predictions, they did not take the further step to enable multi-step prediction and plan generation. Our system can perform multi-step prediction and planning, and as shown in the last stage of development reported in the current paper, it can extract multi-step behaviors from the demonstration that may include multi-objects. Moreover different from other studies [55], we recruited naïve subjects who naturally adapted their demonstrations based on the robot's imitation performance, and used some 'motionese' features enabling the robot to imitate complex actions.

While many studies focus on development of skills corresponding to one particular stage of development, some others modelled long-term development that is achieved in several stages, similar to our work. In their survey of the ontogeny of tool use, Guerin et al. [13] formulates concrete recommendations on general mechanisms of sensorimotor development, and knowledge representation of actions-object relationships. Cangelosi et al. [56] identifies a number of key challenges in developmental robotics, and designs a practical roadmap for developmental robotics which includes a series of milestones such as action learning, language development, social learning,

and cognitive integration. Law et al. [57] realized a staged development with iCub that models an infant from birth to 6 months. iCub, through motor babbling driven with a novelty metric, starts from uncontrolled motor movements, passes through several distinct behavioral stages, and achieves reaching and basic manipulation of objects, similar to the human infants. Our system, on the other hand, develops beyond motor skill formation, and can even engage with humans, observing the demonstrated actions and representing them in terms of robot's sensorimotor and prediction structures learned in the previous stages. A longitudinal development was realized also by Hart and Grupen [58], where a robot self-organizes its sensorimotor space by assembling basic actions into hierarchical programs in a bottom-up way, and by learning to apply these programs in novel contexts in a top-down fashion. The staged learning of behaviors is guided by an intrinsic rewards mechanism that maximizes detection of and acting on affordances with the corresponding behaviors. Hart and Grupen's work is focused on behavior formation in a staged progression with mechanisms similar to accommodation and assimilation [24] through the so-called affordance discovery motivator with emphasis on closed-loop control programs as coupled dynamical systems. In our system, on the other hand, the behavior formation and affordance discovery are decoupled with stronger emphasis on learning prediction abilities based on discovered object affordances using simpler representations in the behavior level. Development of more complex behaviors such as grasping with different strategies requires learning of closed-loop controllers that are guided in real-time by object affordances and tactile feedback as we discuss in the Discussion Section.

A large body of work has been accumulated in the last two decades in this direction with varying degrees of granularity in modeling, fidelity to biological development, and the target level of infant intelligence. However, a coherent integral model that can explain and reproduce various stages of development altogether in a consistent way is still missing[59] for which this work may serve as a contribution towards this direction.

III. DEVELOPMENTAL FRAMEWORK

Our aim is to enable the robot to undergo a developmental progression similar to those of infants. In this section, after providing the initial sensorimotor representations and built-in skills of the robot, we give details of the methods that facilitate such a learning competence. In separate stages, the robot finds behavior primitives which entail similar events (behavior primitive discovery), it learns how these events change the external world (affordances learning), and it deduces how observed behaviors of other agents such as humans can be mapped to own behavior repertoire (imitation learning). Throughout the experiments, only the structures emerging from the described sensorimotor representations are used in subsequent learning stages. While learning new skills, the previously learned skills are kept fixed for simplicity.

A. Built-in perceptual and motor skills

Our staged developmental framework is realized using the robotic system consisting of a 7 DOF Motoman robot arm

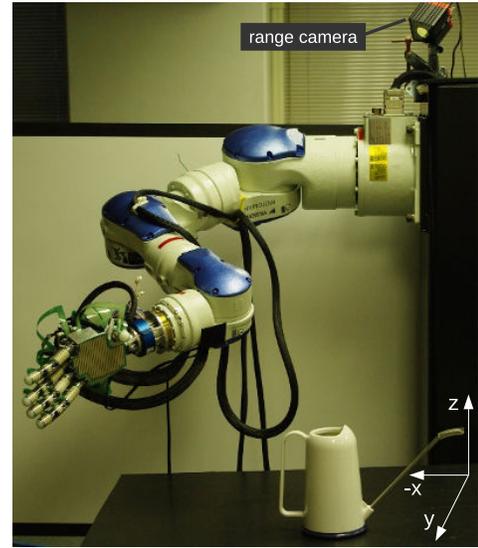


Fig. 2: The hand, arm, range camera, and the environment of the robot.

and an anthropomorphic five fingered 16 DOF Gifu robot hand mounted on it, as shown in Fig. 2. The maximum reach of Motoman arm and Gifu hand are 123 cm. and 23 cm., respectively. There are tactile sensors distributed on the surface of the fingers and palm. For environment perception, an infrared range camera (SR-4000), with 176x144 pixel array, 0.23° angular resolution and 1 cm distance accuracy is used. With the built-in knowledge of object detection and hand-eye coordination, the robot can spot, observe, and manipulate the objects that are within the camera view. In the rest of this section, the details of these built-in perceptual and motor skills of the robot are given.

1) *Tactile perception*: The initially available tactile sensors on the palm of the robot malfunctioned during the experiment. So, we emulated those sensors by comparing the position of the object and the palm and fingers; that were computed using camera and forward kinematics of the robot, respectively. The obtained tactile percept is sampled at 50 Hz., and lumped into 2 binary features: *palm contact* and *finger contact*. The tactile perception observed during action execution (that takes for 4 sec.) is encoded in the following *tactile trajectory* (T_{traj}^r):

$$T_{traj}^r = (p_1, p_2, \dots, p_{200}, f_1, f_2, \dots, f_{200})$$

where r refers to 'raw' encoding. p_t and f_t correspond to the palm and finger sensation at timestep t , respectively; and can be either on (1) or off (0).

Another more compact way of representing tactile perception during action execution is to segment the trajectory to successive 0 and 1's, and represent the tactile trajectory based on duration of successive segments as follows:

$$T_{traj}^c = (d_1^{p=0}, d_1^{p=1}, d_2^{p=0}, \dots, d_n^{p=0}, d_n^{p=1}, d_1^{f=0}, \dots, d_n^{f=0}, d_n^{f=1})$$

where T_{traj}^c refers to compact tactile perception. $d_1^{p=0}$, $d_1^{p=1}$, $d_2^{p=0}$, $d_n^{f=1}$ correspond to the duration (number of timesteps) for the 1st successive no-palm-contacts, for the 1st successive palm-contacts, for the 2nd successive no-palm-contacts, and for

the n^{th} successive finger-contacts, respectively. Thus palm and finger contact durations satisfy the identity: $\sum_{i=1:n} d_i^{p=0} + d_i^{p=1} = 200$.

2) *Visual perception*: The range image obtained from the infrared range camera is used to compute a number of features from the objects in the environment as follows:

Object detection: First of all, 3D positions of the pixels are transformed to the robot coordinate frame (shown in Fig. 2). Next, after the pixels outside the region of interest (table) are filtered out, the remaining pixels of the range image are considered to belong to one or more objects that are segmented by the Connected Component Labeling algorithm [60]. In order to reduce the effect of camera noise, the pixels at the boundary of the object are removed, and the Median and Gaussian filters with 5×5 window sizes are applied. Finally, a feature vector for each object is computed using the positions of the corresponding object pixels as detailed in the next paragraph.

Object features: The object feature vector includes a binary feature for *object-visibility*, and a number of features related to *size*, *position* and *shape* of the object. The points with minimum and maximum values along x, y, and z axes are computed and their difference along these axes are used as three *size*-related features. The 3D position of the object center is used to represent the *position* feature. The last feature channel encodes *shape*-related features, where the distribution of the local surface normal vectors are used. Specifically, histograms of normal vector angles along the latitude and longitude are computed using 20° bins (see [9] for details). Finally, with this encoding, the object feature vector (\mathbf{f}) is composed of $1 + 3 + 3 + (18 + 18) = 43$ features.

$$\mathbf{f} = (\text{visibility}, \text{position}, \text{size}, \text{shape})$$

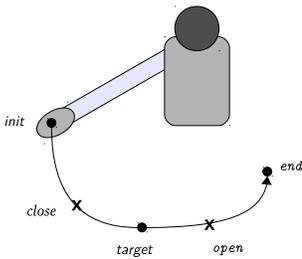


Fig. 3: The parameters of the *swipe* action are illustrated. Any of these five parameters can be set arbitrarily leading to different effects.

3) *Swipe action*: The robot is equipped with a generic *swipe* action which swings the robot's hand towards a detected object. It takes five parameters (see Fig. 3):

$$\text{swipe}(\text{init}, \text{target}, \text{end}, \text{close}, \text{open})$$

The *init*, *target*, and *end* parameters determine the start, middle and end points of a minimum-jerk trajectory. They specify displacement vectors in 3D space with respect to the object

center, where the *target* is typically set to a small value (i.e. small offset from object center) to enable physical interaction. The *close* and *open* parameters encode the flexion and extension of the hand along the trajectory, and take continuous values between $[0-1]$, where 0 and 1 specifies the first and last points of the trajectory, respectively. No activation is specified with a value of -1 . Initially, the hand is assumed to be open. The hand orientation is set such that the wrist is always parallel to the table plane, and the hand is pointing at shoulder-wrist direction. The inverse kinematic method described in [61] is used to compute the joint trajectory using palm center as the end effector. The duration of *swipe* action is fixed to 4 sec.

The execution of the same *swipe* action over the same object with different parameters can produce different effects. Depending on how the *target* parameter is set, the robot can hit the object with its palm, fingers, or may not hit at all. Additionally, depending on the palm-object contact dynamics, the robot may or may not grasp the object. The grasped object can be taken to different final positions or can drop on the way due to opening of the hand. The design of the *swipe* action was based on the goal of having a simple action that, even in this limited experimental setting, will enable the robot to discover meaningful behavior primitives by exploring the parameter space of the action.

4) *Grasp reflex*: With a built-in *grasp-reflex*, if the robot senses a contact at its palm, the fingers are flexed closing the hand. Furthermore, at any moment, fingers can be extended randomly and in this case the robot hand is opened even if there is an object inside.

B. Stage I: Behavior primitive discovery

The robot, through physical interaction with the world and observation of the changes in the environment and self, gradually learns shaping the parameter space in a bottom-up manner. The rest of this section presents the algorithmic details of the first stage of this progressive scaffolding of sensorimotor space, where the robot starts from exploring the parameters of the built-in generic *swipe* action.

The sensorimotor system is constrained in the early ages of human development, and these constraints are gradually lifted during development [62, 63]. As discussed in [64, 65] constrained sensing is very useful to deal with the complexity of input stimuli in absence of necessary perceptual processes, and to reduce the task space for more effective learning. Similar to infant development, the action parameters of the robot are constrained initially, and the constraints are released gradually during the development. In the first stage, we realized this constraint releasing mechanism by allowing the robot to explore the *swipe* parameters as follows:

- *target* position is the main parameter, i.e. released constraint, that is explored in this stage. It is set as an exploration offset ($[0-10]$ cm) from object center in different directions;
- *init* position is fixed to back-right-diagonal of the object from the robot's perspective, and set with (10cm, 10cm, 5cm) displacement from object center;

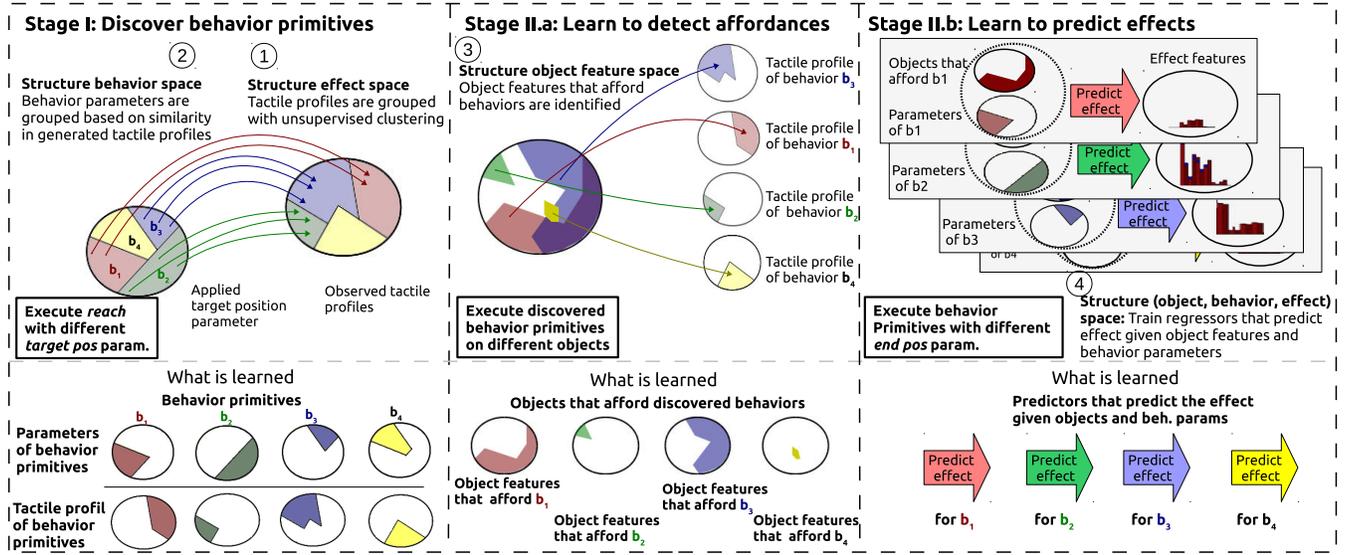


Fig. 4: Illustrative description of Stages I and II. In Stage I, tactile profiles are grouped through unsupervised clustering, and then the corresponding behavior parameter space is obtained based on this grouping. In Stage II, the object feature space is divided to enable affordance detection, and effect prediction is learned. Division of the discs correspond to figurative illustration of structuring feature and parameter spaces. Different colors correspond to structures related to different behavior primitives. Fig. 5 shows how the learned structures are used in Stage III.

- *end* position is randomly set to either *init* position or to a position that is symmetric to *init* along x coordinate with respect to the object center;
- *close* is the time-point that is determined by the object contact and the grasp reflex,
- *open* is the time-point that determines by the relative time of hand opening during hand transport. It is set to -1 if the hand is not to be opened.

$$\text{swipe}(\text{init}, \text{target}, \text{end}, -1, -1) \quad (1)$$

where,

$$\begin{aligned} \text{init} &= (x_{\text{obj}} - 10, y_{\text{obj}} - 10, z_{\text{obj}} + 5) \\ \text{target} &\in \{(x, y, z) : (|x - x_{\text{obj}}|, |y - y_{\text{obj}}|, |z - z_{\text{obj}}|) < (10, 10, 10)\} \\ \text{end} &\in \{\text{init}, \text{init}^{\text{sym}}\} \\ \text{init}^{\text{sym}} &= (x_{\text{init}} + 2(x_{\text{obj}} - x_{\text{init}}), y_{\text{init}}, z_{\text{init}}) \end{aligned}$$

During Stage I, the robot executes *swipe* action N times by selecting the parameters from the allowed ranges as defined above using the uniform random distribution. It observes and stores the readings obtained from tactile experience, gathering N tactile trajectories at the end of the exploration.

Behavior primitives are formed by grouping action executions that generated similar effects in tactile perception. In order to find action execution groups with similar tactile effects, the trajectory experience $\{\mathbf{T}_{\text{traj}}^j\}_{j=1}^N$ is clustered by X-means algorithm where each experienced tactile trajectory is considered as one sample.

$$\{C_i\}_{i=1}^I \leftarrow \text{X-means}(\{\mathbf{T}_{\text{traj}}^j\}_{j=1}^N)$$

where I is the number of clusters.

At the end of clustering, with each tactile trajectory cluster (C_i) a new behavior primitive (b_i) is formed, and associated

to the corresponding cluster. The center of each cluster is assigned a *tactile trajectory signature* corresponding to that primitive:

$$\mathbf{T}_{\text{signature}}^{b_i} = \frac{1}{|C_i|} \sum_{\mathbf{T}_{\text{traj}}^j \in C_i} \mathbf{T}_{\text{traj}}^j \quad (2)$$

where C_i is the cluster corresponding to b_i . $\mathbf{T}_{\text{signature}}^{b_i}$ is the tactile trajectory expected to be generated by the behavior primitive b_i . Signatures will be used to assess the success of the behavior primitive in the later stages.

Next, the parameters of each behavior primitive are computed by taking the average of the executed parameters:

$$[\text{target}, \text{end}, \text{open}, \text{close}]_{b_i} = \frac{1}{|C_i|} \sum_{\mathbf{T}_{\text{traj}}^j \in C_i} [\text{target}, \text{end}, \text{open}, \text{close}]_{\text{swipe}_j} \quad (3)$$

where $[\dots]_{\text{swipe}_j}$ corresponds to the parameters of the *swipe* action executed in trial j , and $[\dots]_{b_i}$ corresponds to the parameters of the new behavior primitive b_i . The set of discovered behavior primitives is denoted as \mathcal{B} .

The left panel in Fig. 4 summarizes the structuring of sensorimotor space in Stage I where the clustering of tactile trajectories is illustrated in ①. The behavior parameters that lead to the corresponding tactile profile clusters are illustrated in ② of Fig. 4. In this figure, the discovered information related to each different behavior primitive is shown with a different color.

At the end of this stage, behavior parameters and tactile trajectory spaces are divided into paired regions to represent the corresponding behavior primitives, and shown in the figurative form in the “What is learned” row.

C. Stage II: Affordance learning

In this stage, the robot executes the discovered behavior primitives on different objects of different positions, orientations and sizes in order to learn the offered affordances. Learning takes place in two steps: First, the robot learns to detect whether a behavior primitive is afforded, given object features. In the second step, it learns to predict how object features would change as a result of an afforded behavior execution, i.e. it learns to predict the continuous effect generated in the objects given the initial object features, and the parameters of the behavior primitive. This is different from our previous formalization [9] where behavior effects were not fixed to success/fail. Here we can safely decide the success of the behavior as behaviors are created based on their tactile profiles in Stage I; thus an execution which does not create such a profile should be regarded as a failure.

Affordance representation: The affordances are represented as triples that consist of the initial percept of the *object*, the *behavior* applied, and the produced *effect* [66]. Recall that *object* feature vector is represented as \mathbf{f} . *Effect* corresponds to the difference between the final and initial perception of the robot, and is defined as the vector difference between the final and initial features: $\mathbf{f}_{\text{effect}}^{b_i} = \mathbf{f}^{b_i} - \mathbf{f}$, where \mathbf{f} and \mathbf{f}^{b_i} represents the feature vector of the object perceived before and after behavior b_i is executed. Here ‘after’ refers to the time-point where there is no change in perception anymore. Thus the affordance relation instance, which represents a sample interaction with the environment, is represented as follows:

$$\{ \langle \mathbf{f}_{\text{effect}}^{b_i}, \mathbf{f}, b_i \rangle \}$$

Stage II.a: Learning to detect affordances: In this step, the robot applies the discovered behavior primitives (Eq. (2)) on different objects that differ in shape, size, and orientation to learn object affordances. Typically, execution of behavior primitive b_i is expected to create a tactile trajectory similar to the corresponding primitive’s *signature* ($\mathbf{T}_{\text{signature}}^{b_i}$), Eq. (2). However, if the execution of b_i primitive at trial t does not create the expected effect, i.e. if the obtained tactile trajectory is more similar to another behavior signature, the result of the execution is regarded as failure:

$$result_{b_i}^t = \begin{cases} 1(\text{success}) & \text{if } b_i = \underset{b_i' \in \mathcal{B}}{\text{argmin}} (\mathbf{T}_{\text{signature}}^{b_i'} - \mathbf{T}_{\text{traj}}^t), \\ 0(\text{failure}) & \text{o.w..} \end{cases}$$

Using this training data, a classifier is trained for each behavior primitive to predict its success given entity feature vector that include object features.

$$\text{aff}^{b_i}(\mathbf{f}) \rightarrow \{0, 1\} \quad (4)$$

where $\text{aff}()$ corresponds to ‘Detect-affordance’ classifier. Specifically, we used a Support Vector Machine (SVM) classifier with Radial Basis Function (RBF) kernel to learn this mapping for each behavior b_i , where \mathbf{f} is given as the input, and the corresponding success/fail as the target category.

Stage II.b: Learning to predict effects: In the second step, the robot learns to predict effects created on objects by further exploring the *end* position parameter of the behavior primitives. It executes the behavior primitives with random *end* positions, stores initial and final features of the objects being interacted, and learns to predict the changes in object features based on the initial features and the *end* parameter. We used Support Vector Machine (SVM) regressors with cross-validation to learn such prediction for each behavior b_i , where (\mathbf{f}, end) is given as the input:

$$\text{eff}^{b_i}(\mathbf{f}, \text{end}) \rightarrow \hat{\mathbf{f}}_{\text{effect}}^{b_i} \quad (5)$$

where $\text{eff}()$ corresponds to ‘Predict-effect’ operator, and $\hat{\mathbf{f}}_{\text{effect}}^{b_i}$ denotes the predicted ($\hat{\cdot}$) effect. This prediction operator can predict the effect given the object feature vector, the index of the behavior primitive, and the *end* position parameter of the behavior.

The middle and right panels in Fig. 4 provide a pictorial description of Stage II. In Stage II.a, learning of the mapping from object feature space to behavior success is illustrated where behavior success is measured based on tactile profiles transferred from the previous stage. Each colored region in object feature space is created by the SVM classifier of a behavior primitive (Eq. (4)), and represents the features that afford the corresponding behavior. On the other hand, the Predict-effect operator is directly illustrated in Stage II.b panel where effect features are represented by bars.

At the end of this stage, object features that afford the discovered behaviors are learned, and the predictors that predict the effects created by these behaviors are trained. These learned structures are shown in figurative forms in the middle and right ‘‘What is learned’’ rows of Fig. 4.

D. Stage III: Imitation

In this stage, in order to teach a complex behavior that involves one or more objects, a tutor displays the behavior in front of the robot. As the robot deals with object manipulation tasks, it focuses only on the trajectory of the objects regardless of the body movements of the tutor. Based on its observation of the object movement, the robot extracts the subgoals from the movement trajectory, finds the actions that produce these subgoals, and executes these actions in sequence to achieve the subgoals and the final goal. Therefore in our system, imitation refers to executing the behavior sequence that achieves the subgoals one by one. This practically corresponds to sequentially moving the object to subgoal positions on the demonstrated trajectory, rather than exactly replicating the demonstration. Note that the detailed imitation trajectory of the object between these subgoals depends on the robot behaviors executed, and might be different from the demonstrated trajectory.

In understanding and achieving subgoals, the affordance perception and behavior primitives that are learned in previous stages are used as basic elements. We already showed that simple goals can be achieved by sequencing these basic elements without utilizing additional mechanisms [9]. If however, the demonstrated actions do not correspond to any robot behavior developed so far, the robot may fail to extract the subgoals

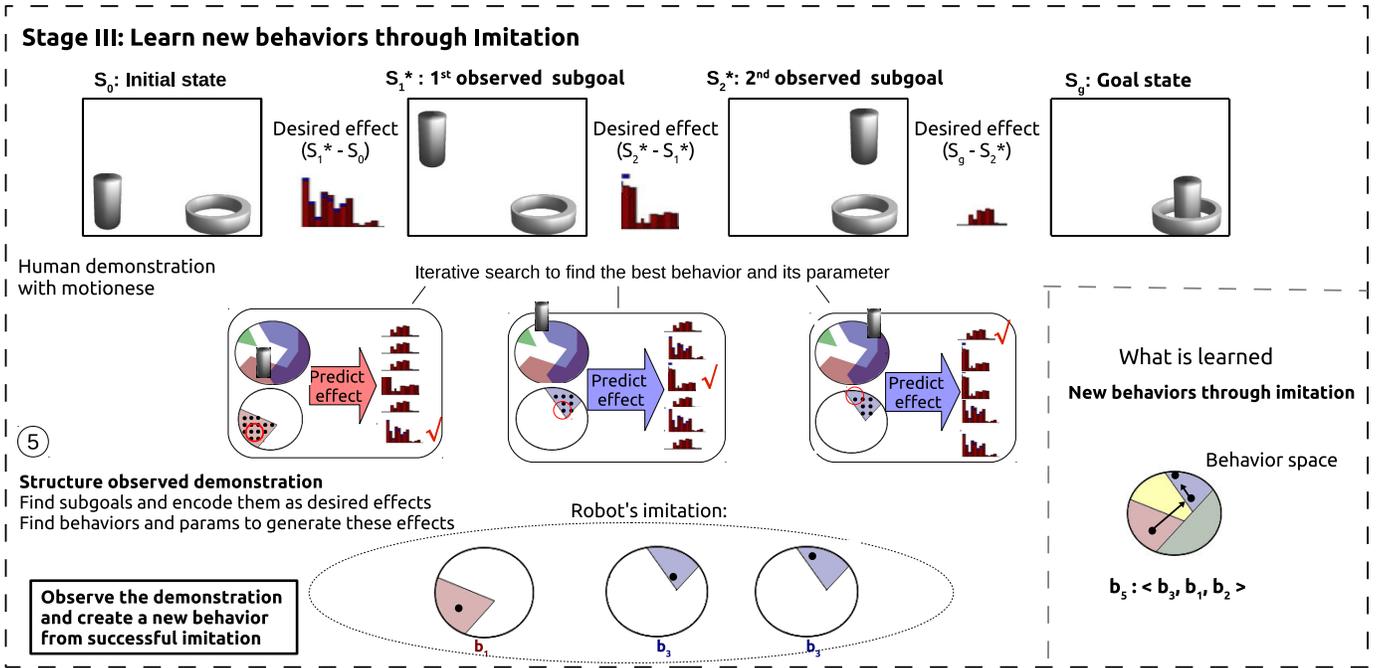


Fig. 5: Illustrative description of Stage III. In this stage, subgoals are extracted from demonstration, and difference between subgoals are encoded as desired effects. The parameters of the afforded behaviors that generate the desired effects are found and sequenced for imitation. Division of the discs correspond to figurative illustration of structuring feature and parameter spaces. Different colors correspond to structures related to different behavior primitives. Formation of these structures in Stages I and II is explained in Fig. 4. The discs filled with different color segments in the middle row for effect prediction correspond to discovered structures in the effect space (Stage II.a in Fig. 4), i.e. each segment illustrates the object features that were learned to afford the corresponding behavior.

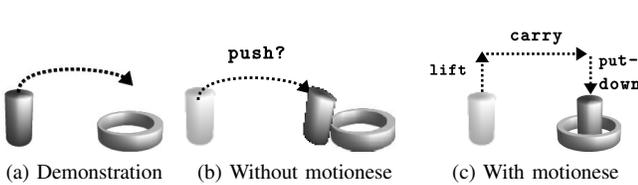


Fig. 6: An example scenario where insertion of a rod into a ring is demonstrated in (a). In order to imitate this task, the robot maps the demonstrated action to its ‘push’ behavior, but fails to achieve the task as the pushed rod would push the ring. However, if the demonstration is exaggerated using motionese, the robot can identify important subgoals, and achieve the task as shown in (c).

and imitate complex demonstrations. For example, when the robot is asked to imitate the action of inserting a rod into a ring as shown in Fig. 6(a), the execution of a behavior that seemingly achieves the goal may fail to satisfy the imitation criteria if the observed trajectory has no interpretation in the sensorimotor space of the robot. In this particular example, if the robot cannot parse the demonstration into executable sub-movements, it may attempt to bring the rod to the observed goal position by simply pushing it to the right, in which case the ring will be pushed away by the rod, rather than the rod being placed in the ring (Fig. 6(b)). On the other hand, when important steps are highlighted through “motionese” as illus-

trated in Fig. 6(c), the robot may extract subgoals represented in its perceptual space, and find a behavior sequence from its behavior repertoire to imitate the action correctly.

The Stage III involves mechanisms to detect important steps of a demonstrated action, which are supposed to be highlighted by a motionese engaged tutor, in order to split a complex action demonstration into doable smaller movement chunks. The details of this motionese based imitation mechanism are provided in the remainder of this section.

Predicting next states: In order to achieve the imitation outlined above, the robot needs to predict the effects of a sequence of behaviors on the objects in the environment, i.e. needs to predict the next state of the environment. State at time t corresponds to the list of feature vectors of the objects, $S_t = [f_{o_0}, f_{o_1}, \dots, f_{o_m}]_t$, where m is the maximum number of objects.

As the behaviors and affordances are learned for single objects in previous stages, the robot assumes that only the features of the corresponding object are affected at a time during the execution of a single behavior. Thus, the next state that is predicted to be obtained by the execution of b_j on object o is computed as follows:

$$\hat{S}_{t+1} = S_t + [\dots, 0, \hat{f}_o^{b_j \text{ effect}}, 0, \dots] \quad (6)$$

where $\hat{f}_o^{b_j \text{ effect}}$ is computed using Eq. (5). Note that as the effect was defined as the vector difference between the final and initial features, the features of object o are updated by

simple summation operator.

Goal emulation: Goal emulation refers to achieving a goal represented as a desired world state ($S_g = [\mathbf{f}_{o_1}, \mathbf{f}_{o_2} \dots]$). If the goal state is achievable in one step, our system can find the best behavior (b^*) along with its parameters (end) using an iterative search as follows

$$\begin{aligned} [\dots, 0, \mathbf{f}_{\text{effect}}^{\text{des}}, 0 \dots] &= S_g - S_t, \\ b^*(\mathbf{f}, \mathbf{f}_{\text{effect}}^{\text{des}}) &= \arg \min_{b_j, end} (\mathbf{f}_{\text{effect}}^{\text{des}} - \hat{\mathbf{f}}_{\text{effect}}^{b_j}) \end{aligned} \quad (7)$$

where $\mathbf{f}_{\text{effect}}^{\text{des}}$ is the desired effect on the object whose features are required to change.

In case a single behavior is not sufficient in reaching the goal state, the robot needs to search for a plan, i.e. a sequence of behaviors, whose total effect is predicted to transform the current state to the goal state ($S_0 \xrightarrow{b_0} S_1 \xrightarrow{b_1} S_2 \dots \rightarrow S_g$). Because prediction is based on vector summation (Eq. (6)), the total effect can be estimated by adding all the predicted effects in sequence to S_0 . In order to find this behavior sequence, a state space search method, namely forward chaining², can be used. A tree of possible next states is formed in this formulation where states are encoded in the nodes of the tree. Each edge corresponds to the prediction of execution of a behavior on an object, and transfers the state to a different state based on the affordance predictors and the summation operators summarized in Eq. (6).

Sequential subgoal emulation: The robot observes the demonstration and extracts the initial and goal states along with the intermediate states (encoded as subgoals) by detecting pauses inserted by a motionese engaged tutor. If no pause can be detected, then a random intermediate state would be picked up as the subgoal state (which may result in a failed imitation attempt).

Once the subgoal states are detected, the robot needs to find the behavior sequence that brings the initial state (S_0) to the goal state (S_g) following the subgoal states. Assuming three pauses were detected along with their states (S_1^*, S_2^*, S_3^*), the robot needs to find four behavior sequences that transfer the initial state to the goal state:

$$S_0 \xrightarrow{\text{beh-seq-1}} S_1^* \xrightarrow{\text{beh-seq-2}} S_2^* \xrightarrow{\text{beh-seq-3}} S_3^* \xrightarrow{\text{beh-seq-4}} S_g \quad (8)$$

Fig. 5 explains how imitation is achieved in Stage III with the structures learned in previous stages. For each observed subgoal (encoded as desired effect in feature space), an iterative search described in Eq. (7) is performed to find the best behavior primitive that generates the desired effect. (In general, multi-step behaviors would be needed to satisfy these transitions, and our framework includes mechanisms to find those behaviors; however, in the experiment we report in this paper, single behaviors were sufficient to provide the the required transitions). Given the object features at time-point t , a grid search is performed in the parameter space of each afforded behavior, predicting the effect on the object, and

the corresponding state transition. Recall that ‘Predict-effect’ ($\text{eff}()$) regressor in Eq. (5) computes the predicted effect for each $((\mathbf{f}_t, end))$ pair where \mathbf{f}_t corresponds to object features in state S_t and end is the parameter searched.

At the end of this stage, the robot learns new behaviors that are encoded as the sequence of behavior primitives and their parameters. Fig. 5, top row shows the extraction of the subgoals from the demonstration. The middle row illustrates the search performed in the behavior space in order to find the behavior primitive along with its end parameter that achieves the same effect for each subgoal. Finally, bottom-most circles show the selected behaviors for each desired effect and their best parameter for the successful imitation.

The role of motionese: Imitation through sequential subgoal emulation is possible only if the subgoals can be detected reliably. We implemented a threshold mechanism based on the speed of the object during demonstration to capture the inserted pauses, and assign a subgoal for each captured object state. The robot finds behavior sequences that are predicted to generate the desired state changes and satisfy subgoals. However, if the demonstrator does not present the subgoals ‘‘correctly’’, the robot may fail imitation as the demonstrator and robot motor spaces are different, and their actions might have different effects on the objects. Thus, we expect the demonstrator to adapt to imitation strategy of the robot by finding ‘‘correct’’ means of demonstration, and changing her teaching accordingly, i.e. by using motionese.

IV. EXPERIMENTAL RESULTS

A. Results of Stage I

The robot performed 64 *swipe* action executions towards a graspable object that is placed in a reachable random position within an area of 20cm \times 10cm on the table. The parameters of the *swipe* action are randomly set according to the allowable ranges detailed in Section III-B. In order to cover the whole range of *target* parameter, the [0–8]cm offsets are divided into grids with 4cm differences, and [0cm – 4cm] further random displacement is set within each grid in each trial. Finally recall that *grasp reflex* and random hand opening³ are also active in this stage. The collected sensorimotor experience is as follows:

$$\{target, end, open, close, \mathbf{f}, \mathbf{f}_{\text{effect}}, \mathbf{T}_{\text{traj}}^c, \mathbf{T}_{\text{traj}}^c\}^{64}$$

Sample effect instances of *swipe* action in different executions are provided in Fig. 7 in the form of object position trajectories, and tactile information. In cases such as trial 2, 11, and 14, the robot hand did not touch the object, but there is a slight change in object position either because of the camera noise or because robot hand is passing over the object and partially occluding it. In some other cases such as trials 23, 29, and 35, the object only touches the robot fingers not triggering grasp-reflex, but being pushed by the fingers. While in trials 38 and 44, the object is grasped and brought to final hand positions; in trials 5, 8, and 26, the grasped object is released with random hand opening.

²We do not imply that human infants necessarily use forward chaining; but rather, we use it as place holder for the corresponding biological mechanisms (e.g. reinforcement learning) we might find in biological systems.

³We simplified the random hand opening effect during grasp, by executing each action with grasp twice, once no-opening of the closed hand and once opening the hand in the middle of *target* and *end* positions.

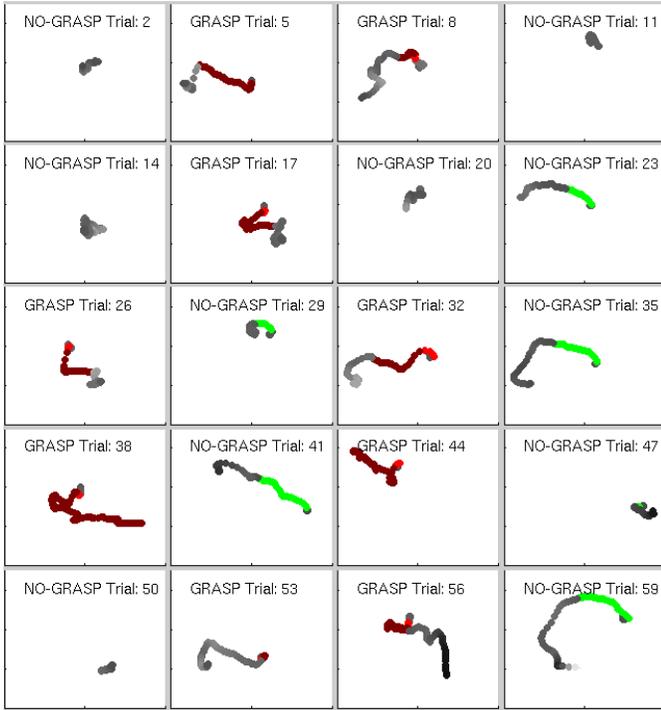


Fig. 7: Top view of sample object position trajectories during exploration of *swipe* movement in Stage I. Gray: No-touch. Bright-red: Palm-only touch. Green: Finger-only touch. Dark-red: Palm+finger touch. See the text for further description. The link to the robot video is provided in Fig. 1.

Although our hypothesis in this paper is that tactile perception plays a major role in distinguishing behaviors for a developing robot, in the experiments we also tested other sensory modalities by applying X-means clustering⁴ in the following channels: changes in object visibility, object position, and tactile perception in raw ($\{T_{traj}^r\}$) and compact ($\{T_{traj}^c\}$) forms. We concluded that combination of palm and finger touch information is sufficient to acquire meaningful behaviors.

1) *Behavior primitives based on object visibility:* Change in object visibility gives the basic information regarding to the existence of the object within robot view during action execution. Four behavior primitives are found when the robot’s execution experience is clustered using:

$$X\text{-means}(\{f_{\text{effect}}(\text{visibility})\})$$

where $f_{\text{effect}}(\text{visibility})$ refers to use of visibility component of the effect feature vector f_{effect} . The results show that each different primitive was characterized by how long the object became invisible during behavior execution. While in one of the behavior primitives the object was always visible, in another primitive, the object becomes invisible in the second half of the executions. Therefore, these two primitives correspond to *no-effect* and *wipe-out* behaviors. However, as

⁴In experiments, we applied both X-means and EM-clustering algorithms which run on real valued and discrete features, respectively; and obtained similar results. So, to be compatible with the rest of the section, we present the results obtained from the X-means algorithm.

the object can become invisible because it is covered by the enclosed robot hand or it is occluded by the moving robot hand, this *wiping-out* may not correspond to pushing object out of the view physically. Additionally, remaining primitives do not have clear and distinct effects. Thus, the physical meaning and distinction of these behavior primitives are ambiguous.

2) *Behavior primitives with similar object position profiles:* The object position change includes more detailed information regarding the visual effect created by that action. Three behavior primitives were found when the robot’s execution experience is clustered using:

$$X\text{-means}(\{f_{\text{effect}}(\text{position})\})$$

where $f_{\text{effect}}(\text{position})$ refers to use of position component of the effect feature vector f_{effect} . The results show that while one of the discovered behavior primitives includes a number of very different action executions, the other primitives include many action executions with no clear distinction. Object position information is a complicated one, and direct clustering in this 3D space does not produce behaviors that are physically distinguishable. Different factors such as partial occlusion, disappearing from view, being lifted or pushed to different sides make an unsupervised clustering challenging.

3) *Behavior primitives with similar tactile profiles:* Tactile sensor gives direct information regarding to the physical interaction of the hand with object, thus tactile profile during action execution can be used to differentiate behaviors that are related to manipulation.

When raw tactile trajectories ($\{T_{traj}^r\}$) are used, 4 behavior primitives are formed in total, two of which correspond to ‘grasp’ actions with different end points. On the other hand, the third primitive represents ‘grasp&release’, and the final one is a mix of no-touch, push and ‘grasp&release’ actions. The reason for construction of 2 classes for one grasp action was probably due to the difference in contact timing of two execution types with different end points.

Fig. 8 shows the discovered 4 behavior primitives when action executions are clustered in compact ($\{T_{traj}^c\}$) tactile trajectory space.

$$X\text{-means}(\{T_{traj}^c\})$$

Each panel corresponds to a different behavior primitive and gives the executed hand trajectories grouped under that behavior primitive. The red and blue markers show time-points where the robot palm and fingers contacted to the object, respectively.

When clustering is performed with compact tactile profiles, qualitatively different and meaningful behavior primitives were obtained. Because compact representation reduces the effect of exact contact timings, the behavior primitives obtained from the compact tactile profiles much better distinguishes the behavior space. The discovered 4 behavior primitives shown in Fig. 8 can be named as

- *push* (temporary touch of finger),
- *no-touch* (no touch),
- *release* (temporary activation of palm and fingers), and
- *grasp* (activation of finger and palm until final position).

TABLE I: The discovered behavior primitives and their parameters.

Name	target	end	open	close
Push	[-0.01, +0.01, -0.10]	[+0.30, +0.07, +0.23]	-1	-1
No-Touch	[-0.05, +0.08, -0.02]	[-0.25, +0.19, -0.15]	-1	-1
Release	[-0.03, +0.03, -0.02]	[+0.04, +0.07, +0.05]	[+0.03, +0.07, -0.01]	[+0.07, +0.07, +0.01]
Grasp	[-0.02, +0.02, -0.01]	[-0.05, +0.06, +0.01]	-1	[+0.05, +0.07, -0.00]

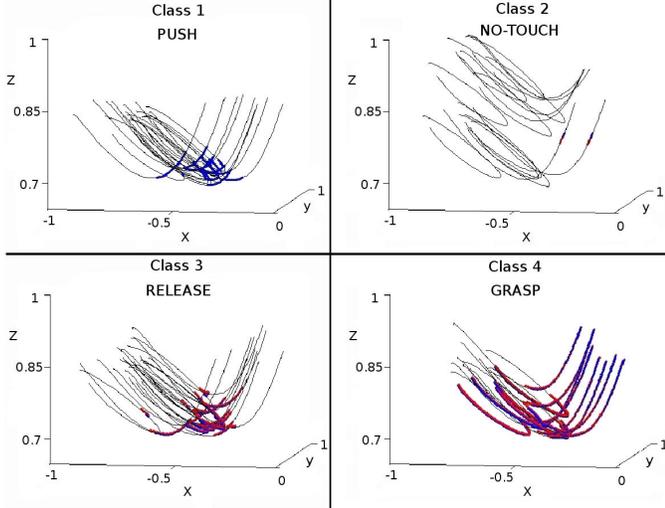


Fig. 8: *Swipe* executions are grouped based on similarity of their compact tactile profiles. Each panel corresponds to a different behavior primitive, and gives the executed hand trajectories grouped under that behavior primitive. Palm and finger touch activation is shown with red and blue marks, respectively.

The parameters of each discovered behavior primitive are computed by taking the average of the parameters in the corresponding cluster (Table I) and transferred to the next stage along with their tactile signatures.

The discovered behavior primitives and parameters that define these primitives have the following characteristics:

- The *target* is the offset from the object-center that determines which part of the robot's hand makes contact with the object. Using this parameter which is unique for each behavior, the robot touches to the object with its palm in *grasp* and *release* behaviors, and with its fingers in *push* behavior.
- *Close* and *open* parameters are the time-points in execution trajectory. The hand clenches into a fist with *grasp* and *release* behaviors when it is close to the object center, and wide-opens with *release* behavior at the end of action execution. *Push* behavior does not change hand-state unless the object is already in the robot's hand. In this case the hand wide-opens in the beginning.
- The *end* position is the offset from the object where the robot brings its hand at the end of the behavior execution. The *start* position is the offset from the object where the robot places its hand prior to interaction. This parameter is fixed and same for all behaviors. If the object to be interacted is already in the robot's hand, the start position

is set as the current position of the robot hand as there is no need to re-position the hand.

As shown in the results (Fig. 8), meaningful behavior primitives were obtained when the tactile trajectory is abstracted from the timing details. One can argue that representing the trajectories in this compact form is an explicit design choice that manipulates the unsupervised behavior discovery process. To confirm the generality of the results, we carried out a cluster analysis in the raw trajectory space using Dynamic Time Warping (DTW), which is a technique that aligns two sequences by warping their temporal profile minimizing the total distance between matching points [67]. DTW provides a metric to compare two trajectories based on their optimal alignment. When we used DTW as the distance metric in clustering raw tactile trajectories, we have seen that the same behavior primitives, namely no-touch, grasp, release and push were obtained. In a different context, [68] also used DTW-like warping in representing the observed action segments that are split based on changes in touch relations.

Discussion: The discovered behavior primitives depend on which clustering methods are used, the parameters of these methods, and the particular feature channels that are being used for clustering. Therefore, different aspects of this clustering process should be analyzed to understand the potentially different ways of behavior formation. In our analysis, we applied the clustering method in different feature spaces, and concluded that meaningful behavior primitives can emerge when the tactile trajectories are used for clustering. In comparing clustering performances of different channels (such as tactile, visibility and position), we used X-means method, which does not require any parameter tuning (such as the target number of clusters). Based the clustering analysis with compact trajectory representation, we concluded that behavior primitives emerge when the executions are grouped based on tactile information which is abstracted from the timing details. We further verified that this conclusion is not simply due to the compact representation we chose but can be obtained via a clustering on raw trajectories by using a distance metric that takes into account temporal shifts (i.e. DTW). More complex and noisy interactions require use of more advanced algorithms such as longest common subsequence (LCSS) [69] or Hidden Markov Models [70].

B. Results of Stage II

This section provides the results obtained in developmental Stage II where the behavior primitives discovered in Stage I are applied to a rich set of objects to learn object affordances. In the previous stage, a search in *target*, *open* and *close* parameters were performed and behavior primitives with along with their parameters were discovered. In this

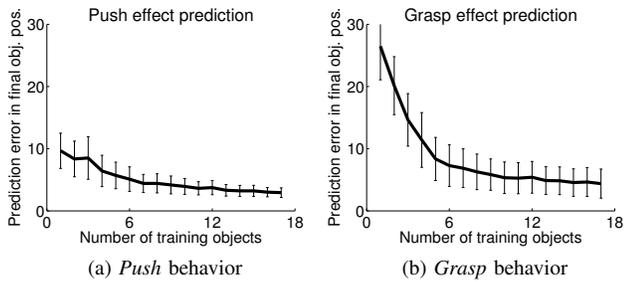


Fig. 12: The error in effect prediction for *push* and *grasp* behaviors with increasing number of training objects. The error corresponds to the distance between actual and predicted final positions. Error bars on prediction errors indicate the mean and standard deviation of the classification results obtained with 50 different training and test sets. Error is given in cm.

With *grasp* behavior, the fingers are enclosed over the object, thus the robot can carry the object to any position defined by the *end* parameter. This is correctly predicted as shown in Fig. 13(a), where predicted final positions of the objects are consistent with the direction and magnitude of the *end* position vector. On the other hand, *push* behavior can only move the object in the front-left direction as the wide-open hand always approaches from the right side of the object, pushing it to the front with the thumb or to the left with the four fingers. Fig. 13(b) shows that the hand-object dynamics during *push* behavior is also learned and correctly predicted. As shown, the movement of the object (the 'x'-'o' vector) is consistent with the *end* parameter (arrow) only if the *end* position is towards front-left direction. For other *end* directions, the object is predicted to move in small amounts with a contact, but this movement is correctly predicted not to be consistent with the *end* direction. Note that if the object was a large or spherical object, it would not afford *grasp* or *push* actions, respectively.

In this section, we showed that the robot can detect various affordances (pushability, rollability, graspability) and can predict the next perceptual state in real world. With these learned mechanisms, given a goal state (directly or through observation), the robot can generate plans which involve parametric behaviors through multi-step prediction. In our previous work [9, 71, 72, 73] we already studied how goal-satisfaction, multi-step planning and goal emulation can be achieved using learned affordances. The generalization performance of the affordance prediction [44], the behavior parameter's effect in prediction and execution [71], and goal emulation through planning [9] were analyzed in detail. Therefore, we do not further analyze the developmental system for planning capabilities. Instead, we show how this system can be further developed to enable complex multi-object imitation with motionese in the next section.

Discussion: While the low-level features are fixed in the current setting, the internal representation of the objects progressively enriches through learning. After learning in Stage II, when encountered with objects, the robot sees them as a set of affordances (e.g. rollable, graspable) along with the low-level initial features. We believe that the progressive organization of

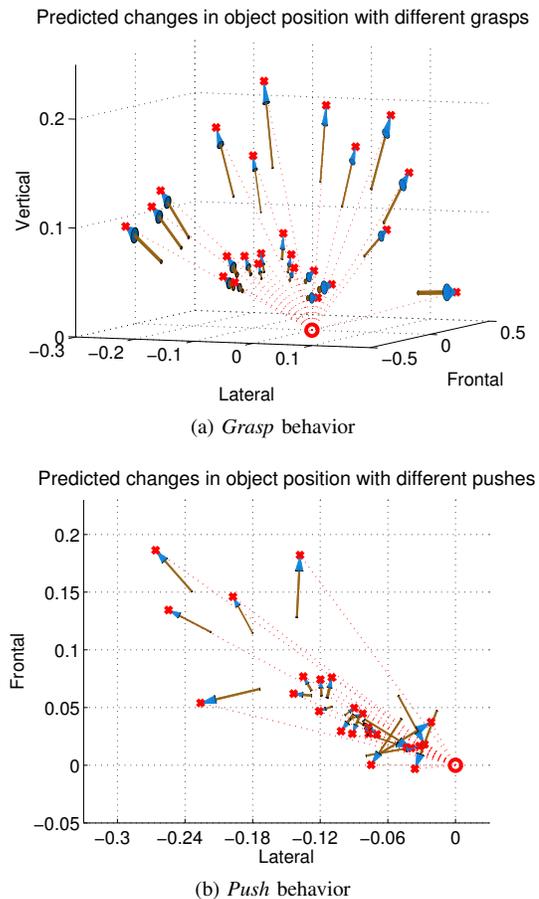


Fig. 13: The robot predicts how object position changes in response to different *grasp* (a) and *push* (b) behavior executions applied to a medium sized box which is graspable but non-rollable. The initial position of the object is shown with the red 'o' marker. Being grasped or pushed in different directions, the object is predicted to be moved to different final positions that are shown with 'x' markers. Behavior's *end* position parameter encodes the direction and magnitude of the robot hand movement after reaching the *target* position, i.e. contacting the object. Thus, each *end* parameter is shown with an arrow in the figure, and placed next to the position that the object is predicted to be moved with that parameter. Note that for the illustration purposes, the magnitude of the arrows are scaled down and they, are shifted to point the predicted final positions of the object. (a) shows the predicted final object positions for the *grasp* behavior where the object can be carried to various positions in the space, in the same direction and proportional to the magnitude of *end* parameter of the behavior. The predictions show that the robot was able to learn that the object can be carried to various locations with the *grasp* behavior. The robot also learned that it can only *push* the object in front-left direction as the hand approaches from right and the hand is not flexed to be afford the pull-back of the object.

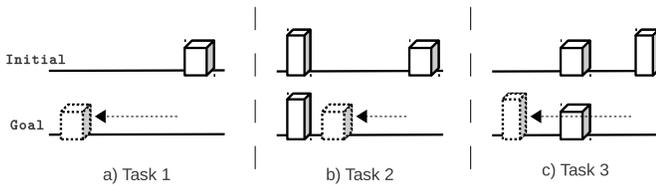


Fig. 14: The tasks used in imitation learning in Stage III.

sensory space based on the action effects provides functionally valid new high-level features, i.e. concepts. Note that the features we used in the beginning were not high-level, and indeed resemble to shape and topological feature detectors that can be found in the primate brain, e.g. orientation and depth selective neurons in the parietal cortex [74, 75].

C. Results of Stage III

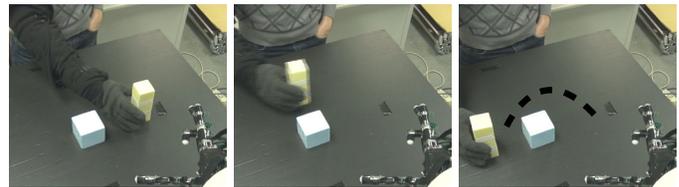
In this section, the experimental results of developmental stage III will be provided where the robot uses the discovered behavior primitives and learned affordances to imitate demonstrations. As described in Section III-D, the robot observes the demonstration and extracts the initial and goal states, as well as the intermediate states (encoded as subgoals) by detecting pauses which may be introduced by a motionese engaged tutor. If no pause can be detected, then an arbitrary intermediate state would be picked up as the subgoal state in the experiments⁶

The aim of this experiment is to investigate whether the proposed imitation and goal emulation framework is intuitive and suitable for naïve tutors in robot teaching setups. For this purpose, 3 subjects who have engineering backgrounds but have no knowledge about this research were recruited as tutors.

We provided the following guideline to the subjects: “The task of the tutor is to teach the robot how to bring the objects to a desired arrangement through demonstration. The tutor starts by showing the initial arrangement of the objects, then performs the same action sequence to bring the objects to the goal arrangement, and finally moves the objects back to their initial positions. He/she is allowed to move only one object a time (similar to one-armed manipulation). The subjects were told that this robot developed itself like a baby by interacting with objects prior to this experiment. We further explained that the robot can push, grasp and release objects, and additionally it can find and execute the behaviors to bring an object to a demonstrated goal position. The robot’s lack of experience with multi-objects was also explained to the tutor.”

We defined three imitation tasks that involve movement of one object in environments with one or two objects. The first task is used to introduce the subjects a typical robot-teaching scenario, where the subjects watch an expert moving one object to the left, and the robot imitating the demonstrated

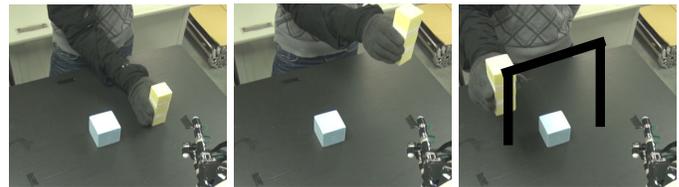
⁶The aim of introducing an intermediate state is to provide feedback to the tutor about the failed imitation attempt of the robot, and provide some indirect hint about the observation mechanisms of the robot. In the current implementation, this state is selected randomly; however this selection can be guided by other type of motionese cues such as extrema of the movement trajectory or by the features extracted from infant robot directed speech.



(a) The first teaching attempt, fast short-path demonstration



(b) Before the last teaching attempt, triangle shaped paused demonstration



(c) The last teaching attempt, rectangular shaped paused demonstration

Fig. 15: Snapshots from naïve tutor’s demonstration to teach Task 3. (a) shows the first trajectory and (b) and (c) show the last two trajectories. These demonstrations took 4, 7 and 9 seconds respectively. As shown, there is significant difference in the speed of demonstration as well as shape and size of trajectories between initial attempt and final attempt of the demonstrator. The trajectories are roughly displayed with dashed and solid lines that correspond to horizontal and vertical movements, respectively. The link to the robot video is provided in Fig. 1.

action (Fig. 14(a)). After familiarizing to the robot-teaching scenario, the subjects were asked to teach Task 2 and Task 3 to the robot. Similar to the previous task, in Task 2, an object is to be moved to the left, but this time next to another object as shown in Fig. 14(b). Finally in Task 3, an object is to be brought to the other side of the other object, as shown in Fig. 14(c). Tasks 1 and 2 have the same complexity as pushing the object to the left satisfies the goal, whereas Task 3 is more complicated as a simple push-left of the object on the right dislodges the other one, failing the objective.

The subjects cleared Task 2 in their first trial by simply moving the object to the left, and the robot imitated this demonstration successfully as there was no ‘obstacle’ on the way. However, Task 3 was challenging so the tutor was required to teach the robot a sequence of behaviors where the object on the right is brought to the other side of the small object while the position of the small object is kept intact. The tutor needed to find out how to teach this behavior to the robot, similar to teaching to a baby, by observing how the robot fails while trying to imitate the tutor’s demonstrations. The tutor attempted to teach Task 3 several times, and the experiment finished when the tutor was able to teach this task.

The tutors performed a number of trials (14, 20 and 18)

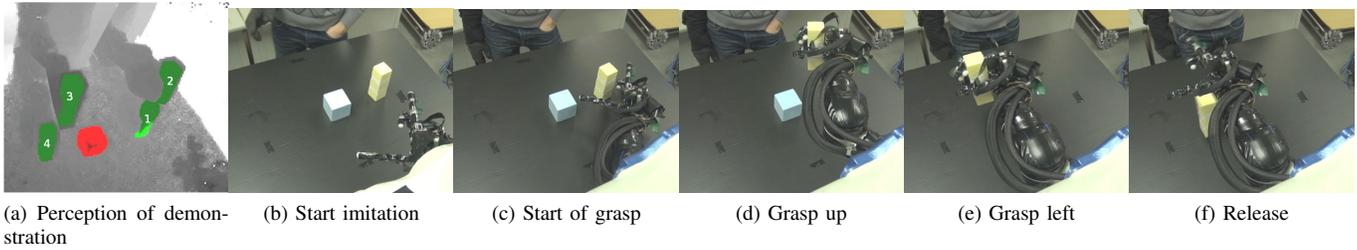


Fig. 16: (a) Snapshots from robot's perception taken during final demonstration of the naïve tutor who learned what type of motionese cues robot uses for imitation. (b-f) The robot's imitation of the square-shaped demonstration observed and segmented in (a).

in total where they started teaching in appropriate way after 9th, 14th and 13th trials, respectively. A number of snapshots from the initial and the final teaching trials of one tutor are provided in Fig. 15. The perception and imitation performance of the robot for the last trial is given in Fig. 16.

The interpretation of tutors' performance in Task 3 is as follows:

- First, the tutor moved the object around the small one in an arc closer to his own body and the table (Fig. 15a). He observed that while the robot was imitating this action, its arm hit the object in the middle and moved it as well. So, he quickly adapted to this failure and did not try this trajectory again.
- As all tutors started with high speeds without any pauses, the robot failed to extract the important points in the trajectories, and to imitate properly. Observing the way the robot failed, after several unsuccessful attempts, all tutors started slowing down their demonstration. Fig. 17 shows how the durations of demonstrations increase by time. While the subjects 2 and 3 tried to find the most compact (and the best timing) for the imitation, the subject 1 exaggerated the duration as well to (probably) see a clear imitation performance.

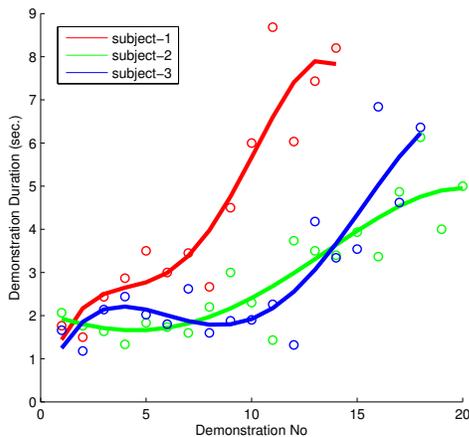


Fig. 17: Evolution of the duration of the demonstrations during teaching attempts for 3 subjects.

- We analyzed one of the tutor's performance (subject 3) in detail. Fig. 18 shows the evolution of the demonstration trajectory. We can clearly see that although the demon-

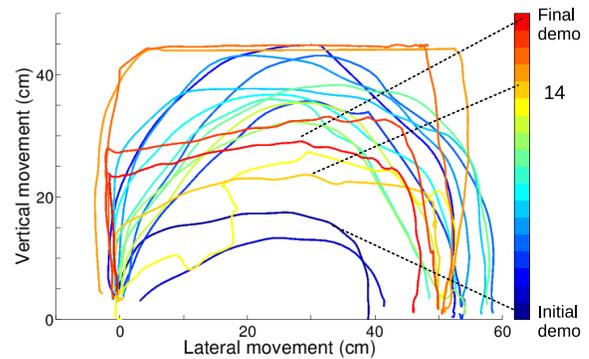


Fig. 18: Evolution of the trajectories that are demonstrated by the subject 3.

stration is small in scale and circular in shape initially, it gradually gets bigger and sharper. In the final two demonstrations, the subject fine-tuned the demonstration by 'checking' the robot's limits (by reducing the size while maintaining the rectangular shape).

- We further quantitatively analyzed how the curvature of the demonstrated trajectories change by time. For this purpose, a curvature value is computed for each point by taking derivative of the slope around the corresponding point. Fig 19 gives the median curvature values of each demonstration. As shown, the trajectory is more curved in the initial demonstrations, and it becomes more straight towards the final demonstrations where the tutor starts being successful in teaching the task. Towards the end, the tutor probably understands that the robot perceives and imitates the demonstrations based on its actions that can only move the objects in straight trajectories. The curvature of the trajectory decreases significantly in demonstration no 14 as shown in Fig. 19. The corresponding trajectory, shown in Fig. 18, suggests that that tutor starts understanding the imitation strategy of the robot.
- Next, we analyzed the pauses inserted by the tutor during his demonstrations. Fig 20 gives the number and the duration of the pauses that are inserted by the tutor during his successive demonstrations. In order to detect the pauses in the very noisy trajectory, the speed trajectory is

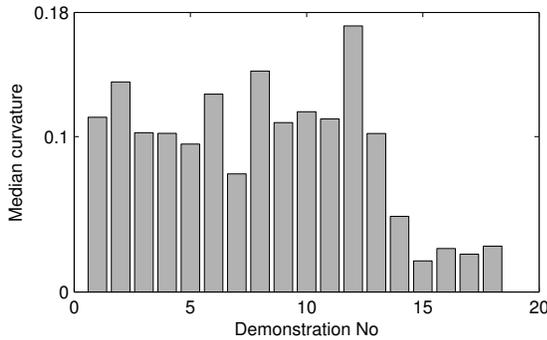


Fig. 19: Evolution of the curvature of the demonstrated trajectories. Each bar corresponds to the median curvature of the demonstrated trajectory where curvature in each point is computed by taking the derivative of the slope (tangent vector) at that point. As shown, demonstration becomes less curved (on average) after demonstration no 14. Please see Fig. 18 for the corresponding trajectories.

computed and smoothed with mean filter. Next, the local minimum points in the speed trajectory are detected and labeled as pauses. As shown, the tutor increased both the duration and the number of pauses, especially after demonstration no 14.

Discussion: In the imitation experiments, it was expected that the subjects would modify their task demonstrations over several trials, however “how” they would modify their demonstrations in terms of shape, speed and amplitude was not trivial to predict. While we designed trajectory segmentation mechanism based on the detected pauses, our results go beyond whether pauses can be exploited by the tutors to teach the robot, and reveal that the demonstrated trajectories change ‘progressively’ in a non-trivial way: The naïve subjects demonstrated the task more slowly (Fig. 17), made bigger and sharper movements (Figs. 18 and 19), and inserted more and longer pauses between movements (Fig. 20) in the later demonstrations. After each demonstration, our robot tried to imitate the observed action using its action primitives. This trial of imitation allowed the subject to recognize how the robot interpreted their demonstrations and what it could generate to reproduce the interpreted actions. Therefore, the evolution of the subjects’ demonstrations indicates how they recognized the robot’s ability. A recent study, which analyzed caregiver’s task demonstration to infants, also showed that non-successful imitation of infants significantly influenced caregivers’ subsequent demonstration. The caregivers emphasized the parts or aspects of their demonstrations depending on infant failures [76]. Our results show a similar effect but provide better insights into the underlying mechanism of task demonstration.

V. GENERAL DISCUSSION

We believe that our study may be advantageous in robotics research in unstructured environments compared to the design-based approaches. In general, robot behaviors are either manually designed or taught via demonstration, with no regard

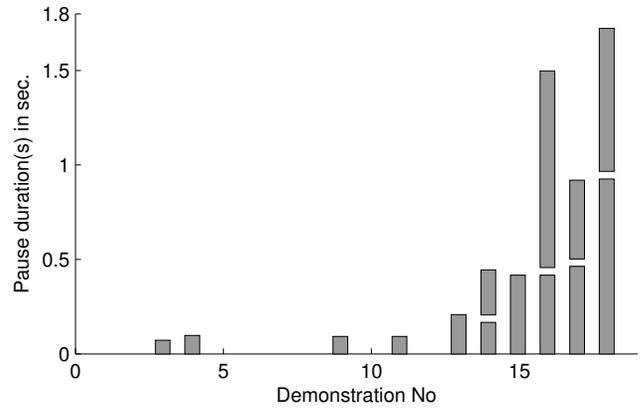


Fig. 20: Evolution of the number and the duration of the pauses inserted in successive demonstrations. Each bar corresponds a pause whose duration is given by the height. The tutor starts inserting two pauses after demonstration number 14. Please see Fig. 18 for the corresponding trajectories.

on how similar/different they are in the perceptual view of the robot. Our study provides a more embodied method for creating a repertoire of primitive behaviors, whose categories are autonomously decided to the extent the robot can differentiate with its own sensors. Classical robotic studies rely on the kinematic/dynamic and sensory models of the robot to infer the consequences of its interaction with the environment. Such model based approaches may fail to capture the details of the dynamic interactions of the robot with the environment due to modeling errors. Our study proposes a data-driven approach towards modeling the interactions of the robot in its own perceptual space.

With such a data-driven approach, we showed that the robot can discover complex sensorimotor skills starting from simple initial perceptual and motor capabilities. The developed skills are grounded within robots initial capabilities and hardware, i.e. sensor and actuation modalities. We can argue that stationary robots with similar tactile and depth perception capabilities, and anthropomorphic hand-arm systems would undergo similar development if the same techniques are used. However, slight changes in these capabilities would generate different set of behavior primitives, affordances, thus different imitation capabilities. For example, a robot that cannot differentiate the tactile difference between palm and fingers, either cannot discover the four behavior primitives, or requires different exploration strategies to form similar primitives. Or if the speed of the hand is very high during the initial swipe action, the objects are pushed away before getting enclosed inside robots hand, therefore grasp and release primitives cannot be discovered. Therefore, despite a number of parameters that are set experimentally (such as the speed of the hand), the progressive use of tactile, visual and social cues are applicable to different manipulator robot systems. We can also argue that while the features and methods used for behavior formation are specific to manipulator robot systems, the principles used in Stage II and Stage III can be applied even to mobile robots, which can learn object affordances and prediction

skills through exploration; and imitate complex actions with sequential subgoal emulation.

While our study provides hints about how sensorimotor development can be achieved in robots, realizing a truly developmental system with life-long learning capabilities still stands as a big challenge. In order to realize such a development in a real robot, we took a number of shortcuts, and initiated the perception and motor system from a state where it is assumed that some learning has taken place. We enumerate a number of assumptions that we made in designing the initial capabilities of our system, and refer to other studies that have investigated the learning of such capabilities.

- First of all, our robot has the built-in capability of detecting objects in the environment. Alternatively, object concept can be learned autonomously through exploration. In [77], a manipulator robot explores the environment with grasp actions, monitors motion of entities, and discovers objectness by relating visual and haptic cues. Similarly, in [78], a mobile robot discovers object instances by tracking spatio-temporal clusters in its sensory experience, and forms object classes based on shapes and effects generated with different actions.
- In our setup, the robot manipulates and observes single objects therefore, we did not need to equip the robot with an attention system (e.g. [79]), which limits our framework to work with single object trajectories. Our system can be integrated with an attention system such as [80], where the robot discovers an attentional landscape that is modulated by its own body and motor programs.
- In our work, we started the developmental progression at the stage where hand-eye coordination and basic reaching capability exist. The basic reaching capability can also be autonomously learned either by acquiring visuomotor maps or by learning the kinematic structure in a developmental way [81, 82, 41, 83].

The experiments indicated that through staged clustering, classification and regression, which are applied to different dimensions of the sensorimotor space, meaningful perceptual and motor categories can be obtained. These categories and the learned predictors can be further used to bridge the gap between self-discovered structures and demonstrated tasks. By using generic non-parametric methods such as X-means and SVMs, we mitigated the emphasis placed on particular methods and parameters, and we focused on how to explore the vast search space. While we explicitly designed the means of search (clustering in Stage I, classification and regression in Stage II, segmentation in Stage III), in real biological systems the search of exploration space is guided by biological constraints of the self (e.g. infants start with limited visual perception, or cannot explore the whole environment because of the limitations in their locomotion capabilities), or scaffolded by the parents. In robot development settings, what/how to explore next can be further guided by Intrinsic Motivation [84, 58], which we currently study for emergence of development order of simple-to-complex affordances in [85].

VI. FUTURE WORK

The biological systems learn different skills simultaneously, and in a life-long ongoing process. As we discussed in Section II, infants focus on learning of particular skills in different stages of development, where the transition between the stages emerges automatically in a seamless manner [13], and is probably a natural consequence of changes in representations [86]. It is also suggested that humans start from a reduced set of degrees of freedoms in the initial stages, and later gradually lift all restrictions [87]. Our system, on the other hand, simplifies skill development by separately learning different skills in different stages, and freezing the development of the previously learned skills while learning new ones. The staged approach (specifically freezing the results of a stage before moving on the next one) was a choice, due to the simplification of the construction and analysis of competences acquired at the end of each stage. Although using on-line and incremental machine learning methods would allow the robot to keep all stages plastic, the stability of the whole system as well as the interaction between the simultaneous activities of these stages posed a greater challenge that went beyond the context of this study. The limitation of freezing skill development can be relaxed by progressively activating skill learning modules, and letting the robot to continue developing all the skills in parallel in the future implementations. Having said this, [88] suggests that a strategy, which adaptively alternates freeing and freezing degrees of freedom, can cope better with environmental perturbations. We also believe that an approach that adaptively alternates skill development can cope better with the large spaces of learning and exploration. In the future, we plan to study how the proposed approach can scale up to support complex tasks that include manipulation and interaction of multiple objects with the increasing multiplicity and diversity of real-world datasets [89]. Mechanisms like alternating skill development [88], structural bootstrapping [90, 91, 92], transfer learning [93], and intrinsic motivation [94, 85] should be utilized to cope with the complexity of the learning in such large sensorimotor spaces.

The challenge the robot faces for learning complex skills such as inserting one object into another is two-fold. It needs to learn both action related properties (lift, move, drop), and object related properties (e.g. the object below should be concave and its hole should be larger than the object being dropped). In learning action related properties, we discuss that even the robots which only know about single-object affordances can acquire multi-object actions through imitation that is coupled with tutor's scaffolding. After the robot obtained the basic strategy (lift, move, drop), it should further explore the action parameters (height of lifting action, movement trajectory depending on relative size and shapes of objects, force feedback during insertion, etc.) and adapt the new 'insert' primitive accordingly as we discuss in the next paragraph. Regarding the second problem, i.e. learning of 'relational affordances', we recently showed that the robot can benefit from a hierarchical structure where pre-learned basic affordances are (re-)used as inputs of complex affordance predictors, bootstrapping the learning of complex affordances

[92]. In the same study, we also showed that (re-)using pre-learned basic affordances in active selection of objects to explore next also speeds up learning of complex affordances. Moldovan and Raedt [95] also studied learning of relational affordances, re-using pre-learned affordance models, where the focus was on modelling the spatial configuration of objects with generative methods in a probabilistic relational setting. Our work employs discriminative approaches to make sense of the sensorimotor world, however as we discussed in Section II-B, generative methods have also significant potential to capture the underlying structure of the world. It is left as a future work to address to what extent these two approaches should coexist in a developing cognitive system, or whether one has definite advantage over the other and thus only one should prevail.

At the motor side, our object-based imitation, which uses learned effect prediction and sub-goal emulation mechanisms, resides between full trajectory level imitation and goal emulation. The demonstrated behavior is assumed to be learned if the robot successfully achieves the goal. In the robot's world, the learned new behavior corresponds to the behavior sequence, where parameters of each behavior are set based on the initial/final position of the object being acted. This learning corresponds to acquisition of one instance of a behavior class, and requires further exploration with additional parameters. In our example scenario, the robot learns to move one object to the other side of another object. However, the learned behavior representation does not include any parameter related to the object that acts as an obstacle in our case; and the robot cannot successfully execute the learned action if the height or position of the obstacle is different from the learned configuration. Therefore, we require stronger behavior representations for complex behaviors, in order to learn coupling parameters of motor control with object features in a flexible and natural way. To exhibit the full-range of imitation and movement capabilities that can be expected from a general purposes robotic system, our framework should incorporate more flexible motion representation frameworks such as Dynamic Motor Primitives (DMPs) [96, 97], and perform learning with these representations. One approach might be to learn DMPs from robot's own action trajectories, and adapt DMP parameters based on object features. Stulp et al. [98], for example, uses Reinforcement Learning to adapt DMP parameters while learning from sequences of motion primitives. Our robot that observes several demonstrations of different actions (such as inserting-into-container or moving-over-obstacle) should also be equipped with such chunking [98] and abstraction [58] mechanisms, in order to develop higher-level of conceptual knowledge.

Finally, our system detects pauses in the demonstration trajectory and finds the sub-goals based on the detected pauses. We set the fixed pause threshold to one second and the naïve tutors were able to learn in this setting; however more experiments with different thresholds and a more thorough analysis are required to exactly understand the role of pause and how it effects the learning curve of the naïve subject. Motionese displayed by the real caregivers, on the other hand, is accompanied by additional and richer set of scaffolding

signals including social signals such as gaze and speech. In our robotic experiments, we did not utilize such social and gaze signals, and possibly due to this, adaptation of the naïve tutors to the imitation mechanism of the robot was slow. It is a future work to augment the sub-goal detection system with other types of motionese and social signals such as gaze and speech.

VII. CONCLUSION

In this study, we have realized a staged developmental system, where the robot achieves higher level cognitive skills by organizing its continuous sensorimotor space incrementally. We focused on three levels of skill acquisition in robots, following a developmental timeline, similar to those of infants'. First, we studied how a robot can discover meaningful discrete primitives by self-organizing its continuous behavior parameter space. Next, we studied how further exploration with the discovered behavior primitives can lead the robot to learn object affordances and associated predictors that can quantitatively anticipate the effects that can be created with these primitives. Finally, we studied how discovered behaviors and learned affordances can be used in bootstrapping the imitation system of the robot. We showed that affordance based goal-emulation ability, together with motionese, enables the robot to imitate demonstrated complex actions that are not directly represented in its behavior repertoire.

The main contributions of this paper can be summarized as follows:

- From the robotics point of view, this study realized staged learning of a wide range of skills (behavior primitives, affordances, emulation, imitation) in **an integrated framework**. Throughout the development, the structures emerged from the previous sensorimotor experiences are directly used as the sole building blocks in the subsequent stages, facilitating the development of demonstrated complex cognitive capabilities.
- Learning a large variety of skills through interaction **in the real world** is a difficult challenge taking into account the fragility of the robotic systems. Different from our previous work, and many other developmental robotics studies summarized in Section II-B, development of the complete system was achieved in the real world on the physical hardware.
- In our robotic system, we showed that the use of **simple-to-complex** perceptual skills, i.e. first tactile, then visual and finally social cues, was necessary and sufficient for the progressive development of the targeted sensorimotor skills. This is in accordance with the characteristics of infant development [62], and thus we may speculate that the mechanisms we used in our implementation can be considered as a possible model for those mechanisms in the human infant
- We showed that while the necessary mechanisms for goal emulation can be acquired by self-interaction, in order to bootstrap learning of more complex behaviors in a feasible time, the robot needed to **leave self-exploration** strategy and **engage in observational learning** by interacting with tutors. This developmental order is also

consistent with infant development as we discussed in detail in Section II-A.

- In human infants, the development of goal-emulation ability precedes the skill of object-directed imitation; but the underlying mechanism is widely unknown. In this work, we utilized goal-emulation ability for generating chunks of movements that are sequentially chained to bootstrap imitation capability on our robot. This robotic solution can be seen as a prediction for the **mechanism of imitation development in infants**.
- We showed that motionese can be used to bridge the gap between the interacting agents with different movement capabilities, such as the human tutors and the arm-hand robot we employed in this study. Furthermore, our experimental data indicated that **naïve tutors** who are not informed about the imitation mechanisms of the robot, changed their teaching strategy, and started displaying motionese even in the absence of explicit feedback from the robot.

While our study provides hints about how development can be achieved in artificial embodied agents towards forming symbols, realizing a truly developmental system with life-long learning capabilities still stands as a big challenge. The robots need to learn progressively higher level concepts that are suitable for high-level reasoning and planning, which could be based on symbols formed via a developmental progression we realized. These concepts should be transferred to other domains and re-used in bootstrapping learning of other concepts. In our work, the sensorimotor space of the robot is organized to enable simple manipulation (and understanding) of the environment in single-object settings. The robot acquired conceptual knowledge by discovering action categories like ‘push’ and ‘carry’ and object categories like ‘rollable’, ‘pushable’. On the other hand, with richer set of behavior primitives and exploration environment, we believe that progressive learning of such knowledge can enable the robot to acquire more complex high-level concepts and other operations such as mental rotation following the same methodology. In our recent work [99], the robot develops other concepts such as ‘unstable’, ‘insertable’, ‘stackable’ by exploring the effects of a richer set of behavior primitives. It discovers high-level object categories, effect categories and logical rules, that are used to encode world state and domain description, later enabling symbolic planning. Current work is underway for modifying the proposed approach to support complex tasks that include manipulation and interaction with multiple objects leading to complex and rich symbol formation that can support development of further cognitive abilities.

ACKNOWLEDGEMENTS

This research was partially supported by European Community’s Seventh Framework Programme FP7/2007-2013 (Specific Programme Cooperation, Theme 3, Information and Communication Technologies) under grant agreement no. 270273, Xperience; by European Community’s Seventh Framework Programme FP7/2007-2013 under the grant agreement no. 321700, Converge; and by JSPS/MEXT Grants-in-Aid for Scientific Research (Research Project Number:

24000012, 24119003). It was also supported in part by a contract with the Ministry of Internal Affairs and Communications, Japan, entitled, ‘Novel and innovative R&D making use of brain structures’.

REFERENCES

- [1] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, “Cognitive developmental robotics: a survey,” *IEEE Tran. Auton. Mental Dev.*, vol. 1-1, 2009.
- [2] A. Stoytchev, “Some basic principles of developmental robotics,” *IEEE Transactions on Autonomous Mental Development*, pp. 122–130, 2009.
- [3] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen, “Autonomous mental development by robots and animals,” *Science*, vol. 291, pp. 599–600, 2001.
- [4] J. Zlatev and C. Balkenius, “Introduction: Why epigenetic robotics?” in *Proceedings of the 1st International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies, 85., 2001, pp. 61–67.
- [5] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, “Developmental robotics: A survey,” *Connection Science*, vol. 15, no. 4, pp. 151–190, Dec. 2003.
- [6] M. A. Arbib, “Perceptual structures and distributed motor control,” *Comprehensive Physiology*, 1981.
- [7] M. A. Arbib, P. Erdi, and J. Szentagothai, *Neural organization: Structure, function, and dynamics*. MIT press, 1998.
- [8] B. Vollmer and H. Forssberg, “Development of grasping and object manipulation,” in *Sensorimotor Control of Grasping: Physiology and Pathophysiology*. Cambridge University Press, 2009.
- [9] E. Ugur, E. Oztop, and E. Sahin, “Goal emulation and planning in perceptual space using learned affordances,” *Robotics and Autonomous Systems*, vol. 59, no. 7–8, pp. 580–595, 2011.
- [10] R. Brand, D. Baldwin, and L. Ashburn, “Evidence for ‘motionese’: Modifications in mothers infant-directed action,” *Developmental Science*, vol. 5, pp. 72–83, 2002.
- [11] E. Ugur, Y. Nagai, H. Celikkanat, and E. Oztop, “Parental scaffolding as a bootstrapping mechanism for learning grasp affordances and imitation skills,” *Robotica*, 2014, published online <http://dx.doi.org/10.1017/S0263574714002148>.
- [12] D. K. James, “Fetal learning: A critical review,” *Infant and child development*, vol. 19, pp. 45–54, 2010.
- [13] F. Guerin, N. Kruger, and D. Kraft, “A survey of the ontogeny of tool use: from sensorimotor experience to planning,” *IEEE Transactions on Autonomous Mental Development*, vol. 5, no. 1, 2013.
- [14] J. Piaget and B. Inhelder, *The Psychology of the Child*. New York, USA: Basic Books, 1966.
- [15] D. A. Rosenbaum, *Human Motor Control*. London, UK: Academic Press, 1991.
- [16] J. G. Bremner, *Infancy*. Malden, MA, USA: Blackwell Publishing, 1988, 2nd edition, 1994.
- [17] J. J. Gibson, *The Ecological Approach to Visual Perception*. Lawrence Erlbaum Associates, 1986.
- [18] P. Willatts, “The stage IV infant’s solution of problems requiring the use of supports,” *Infant Behavior and Development*, vol. 7, pp. 125–134, 1984.
- [19] B. Elsner and B. Hommel, “Effect anticipation and action control,” *Journal of Experimental Psychology*, vol. 27, pp. 229–240, 2003.
- [20] J. Provasi, C. D. Dubon, and H. Bloch, “Do 9- and 12-month-olds learn means-ends relation by observing?” *Infant Behavior and Development*, vol. 24, no. 2, pp. 195–213, Feb. 2001.
- [21] M. Carpenter, K. Nagell, and M. Tomasello, “Social cognition, joint attention, and communicative competence from 9 to 15 months of age,” *Monographs of the Society for Research in Child Development*, vol. 63, no. 4, 1998.
- [22] B. Elsner and G. Aschersleben, “Do I get what you get? Learning about the effects of self-performed and observed actions in infancy,” *Consciousness and Cognition*, vol. 12, pp. 732–751, 2003.
- [23] S. C. Want and P. L. Harris, “How do children ape? Applying concepts from the study of non-human primates to the developmental study of imitation in children,” *Developmental Science*, vol. 5, pp. 1–13, 2002.
- [24] J. Piaget, *The Origins of Intelligence in Children*. New York, USA: International University Press, 1952.
- [25] M. Tomasello, *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press, 1999.

- [26] P. Willatts, "Development of means-end behavior in young infants: Pulling a support to retrieve a distant object," *Developmental Psychology*, vol. 35, pp. 651–666, 1999.
- [27] B. Elsner, "Infants' imitation of goal-directed actions: the role of movements and action effects," *Acta psychologica*, vol. 124, no. 1, pp. 44–59, 2007.
- [28] C. Huang and T. Charman, "Gradations of emulation learning in infants: imitation of actions on objects," *Journal of Experimental Psychology*, vol. 92, no. 3, pp. 276–302, 2005.
- [29] M. Paulus, S. Hunnius, M. Visser, and H. Bekkering, "Imitation in infancy: Rational or motor resonance?" *Child development*, vol. 82, no. 4, pp. 1047–1057, 2011.
- [30] M. van Elk, H. T. van Schie, S. Hunnius, C. Vesper, and H. Bekkering, "You'll never crawl alone: Neurophysiological evidence for experience-dependent motor resonance in infancy," *Neuroimage*, vol. 43, no. 4, pp. 808–814, 2008.
- [31] R. J. Brand, W. L. Shallcross, M. G. Sabatos, and K. P. Massie, "Fine-Grained Analysis of Motionese: Eye Gaze, Object Exchanges, and Action Units in Infant-Versus Adult-Directed Action," *Infancy*, vol. 11, no. 2, pp. 203–214, 2007.
- [32] K. J. Rohlfing, J. Fritsch, B. Wrede, and T. Jungmann, "How can multimodal cues from child-directed interaction reduce learning complexity in robots?" *Advanced Robotics*, vol. 20, no. 10, pp. 1183–1199, 2006.
- [33] Y. Nagai and K. J. Rohlfing, "Computational Analysis of Motionese Toward Scaffolding Robot Action Learning," *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 1, pp. 44–54, 2009.
- [34] Y. Nagai, "From Bottom-Up Visual Attention to Robot Action Learning," in *Proc. of the 8th IEEE International Conference on Development and Learning*, 2009.
- [35] Y. Nagai, A. Nakatani, and M. Asada, "How a robots attention shapes the way people teach," in *Proc. of the 10th International Conference on Epigenetic Robotics*, 2010, pp. 81–88.
- [36] A. N. Meltzoff and M. K. Moore, "Imitation of facial and manual gestures by human neonates," *Science*, vol. 198, pp. 75–80, 1977.
- [37] K. A. Bard, "Neonatal imitation in chimpanzees (pan troglodytes) tested with two paradigms," *Animal Cognition*, vol. 10, no. 2, pp. 233–242, 2007.
- [38] E. Oztop, N. S. Bradley, and M. A. Arbib, "Infant grasp learning: A computational model," *Experimental Brain Research*, vol. 158, pp. 1354–1361, 2004.
- [39] J. M. Romano, K. Hsiao, G. Niemeyer, S. Chitta, and K. J. Kuchenbecker, "Human-inspired robotic grasp control with tactile sensing," *Trans. Rob.*, vol. 27, no. 6, pp. 1067–1079, 2011.
- [40] P. Savastano and S. Nolfi, "A robotic model for reaching and grasping development," *IEEE Transaction on Autonomous Mental Development*, vol. PP, no. 99, pp. 1–12, 2013.
- [41] R. Saegusa, G. Metta, G. Sandini, and L. Natale, "Developmental perception of the self and action," *IEEE Transaction on Neural Networks and Learning Systems*, vol. 24, no. 10, pp. 1–20, 2013.
- [42] J. Tani and M. Ito, "Self-Organization of Behavioral Primitives as Multiple Attractor Dynamics : A Robot Experiment," *Advanced Robotics*, vol. 33, no. 4, pp. 481–488, 2003.
- [43] S. Nishide, T. Ogata, J. Tani, K. Komatani, and H. G. Okuno, "Autonomous Motion Generation Based on Reliable Predictability," *Journal of Robotics and Mechatronics*, vol. 21, no. 4, 2009.
- [44] E. Ugur and E. Şahin, "Traversability: A case study for learning and perceiving affordances in robots," *Adaptive Behavior*, vol. 18, no. 3–4, 2010.
- [45] L. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, "Learning object affordances: From sensory-motor maps to imitation," *IEEE Transactions on Robotics*, vol. 24, no. 1, pp. 15–26, 2008.
- [46] F. Wörgötter, A. Agostini, N. Krüger, N. Shylo, and B. Porr, "Cognitive agents: A procedural perspective relying on the predictability of Object-Action-Complexes OACs," *Robotics and Autonomous Systems*, vol. 57, no. 4, pp. 420–432, Apr. 2009.
- [47] J. Sinapov, C. Schenck, K. Staley, and A. Stoytchev, "Grounding semantic categories in behavioral interactions: Experiments with 100 objects," *Robotics and Autonomous Systems*, vol. 62, no. 5, 2014.
- [48] S. Griffith, J. Sinapov, V. Sukhoy, and A. Stoytchev, "A behavior-grounded approach to forming object categories: Separating containers from noncontainers," *IEEE Transactions on Autonomous Mental Development*, vol. 4, pp. 54–69, 2012.
- [49] S. Ivaldi, S. Nguyen, N. Lyubova, A. Droniou, V. Padois, D. Filliat, P.-Y. Oudeyer, and O. Sigaud, "Object learning through active exploration," *IEEE Transactions on Autonomous Mental Development*, 2013, accepted.
- [50] O. Yuruten, E. Sahin, and S. Kalkan, "The learning of adjectives and nouns from affordance and appearance features," *Adaptive Behavior*, vol. 21, no. 6, pp. 437–451, 2014.
- [51] S. Kalkan, N. Dag, O. Yuruten, A. M. Borghi, and E. Sahin, "Verb concepts from affordances," *Interaction Studies*, 2014, in press.
- [52] H. S. Koppula, R. Gupta, and A. Saxena, "Learning human activities and object affordances from rgb-d videos," *International Journal of Robotics Research (IJRR)*, vol. 32, no. 8, pp. 951–970, 2013.
- [53] S. M. Nguyen and P.-Y. Oudeyer, "Active choice of teachers, learning strategies and goals for a socially guided intrinsic motivation," *Paladyn Journal of Behavioural Robotics*, vol. 3, no. 3, pp. 136–146, 2012.
- [54] M. Lopes, F. Melo, and L. Montesano, "Affordance-based imitation learning in robot," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 1015–1021.
- [55] K. Mochizuki, H. Nobuta, S. Nishide, H. G. Okuno, and T. Ogata, "Developmental human-robot imitation learning with phased structuring in neuro dynamical system," in *IROS2012 Workshop on Cognitive Neuroscience Robotics*, 2012.
- [56] A. Cangelosi, G. Metta, G. Sagerer, S. Nolfi, C. Nehaniv, K. Fischer, J. Tani, T. Belpaeme, G. Sandini, F. Nori *et al.*, "Integration of action and language knowledge: A roadmap for developmental robotics," *Autonomous Mental Development, IEEE Transactions on*, vol. 2, no. 3, pp. 167–195, 2010.
- [57] J. Law, P. Shaw, K. Earland, M. Sheldon, and M. Lee, "A psychology based approach for longitudinal development in cognitive robotics," *Frontiers in neurobotics*, vol. 8, 2014.
- [58] S. Hart and R. Grupen, "Learning generalizable control programs," *Autonomous Mental Development, IEEE Transactions on*, vol. 3, no. 3, pp. 216–231, 2011.
- [59] E. Ugur, Y. Nagai, E. Oztop, and M. Asada, Eds., *Proceedings of Humanoids 2012 Workshop on Developmental Robotics: 'Can developmental robotics yield human-like cognitive abilities?'*, 2012.
- [60] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision, Volume 1*. Addison-Wesley, 1992.
- [61] B. Moore and E. Oztop, "Redundancy parametrization for flexible motion control," in *ASME IDETC*, 2010.
- [62] J. Rutkowska, "Scaling up sensorimotor systems: Constraints from human infancy," *Adaptive Behavior*, vol. 2, pp. 249–373, 1994.
- [63] K. M. Newell, "Constraints on the development of coordination," *Motor development in children: Aspects of coordination and control*, vol. 34, pp. 341–360, 1986.
- [64] M. H. Lee, Q. Meng, and F. Chao, "Developmental learning for autonomous robots," *Robotics and Autonomous Systems*, vol. 55, no. 9, pp. 750–759, 2007.
- [65] J. Law, M. Lee, M. Hlse, and A. Tomassetti, "The infant development timeline and its application to robot shaping," *Adaptive Behavior*, vol. 19, no. 5, pp. 335–358, 2011.
- [66] E. Şahin, M. Çakmak, M. R. Doğar, E. Ugur, and G. Üçoluk, "To afford or not to afford: A new formalization of affordances toward affordance-based robot control," *Adaptive Behavior*, vol. 15, no. 4, pp. 447–472, 2007.
- [67] M. Müller, "Dynamic time warping," *Information retrieval for music and motion*, pp. 69–84, 2007.
- [68] R. Vuga, E. E. Aksoy, F. Worgotter, and A. Ude, "Probabilistic semantic models for manipulation action representation and extraction," *Robotics and Autonomous Systems*, 2014, in press.
- [69] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multi-dimensional trajectories," in *Data Engineering, 2002. Proceedings. 18th International Conference on*. IEEE, 2002, pp. 673–684.
- [70] B. Morris and M. Trivedi, "Learning trajectory patterns by clustering: Experimental studies and comparative evaluation," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 312–319.
- [71] E. Ugur, E. Oztop, and E. Sahin, "Going beyond the perception of affordances: Learning how to actualize them through behavioral parameters," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 4768–4773.
- [72] E. Ugur, E. Sahin, and E. Oztop, "Affordance learning from range data for multi-step planning," in *Proceedings of the 9th International Conference on Epigenetic Robotics*, 2009, pp. 177–184.
- [73] E. Ugur, M. R. Doğar, M. Çakmak, and E. Şahin, "The learning and use of traversability affordance using range images on a mobile robot," in *Proceedings of ICRA'07*, 2007, pp. 1721–1726.
- [74] A. Murata, V. Gallese, G. Luppino, M. Kaseda, and H. Sakata, "Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal are AIP," *Journal of Neurophysiology*, vol. 83, no. 5, pp. 2580–2601, 2000.

- [75] H. Sakata, K. I. Tsutsui, and M. Taira, "Toward an understanding of the neural processing for 3d shape perception," *Neuropsychologia*, vol. 43, pp. 151–161, 2005.
- [76] H. Fukuyama, S. Qin, Y. Kanakogi, Y. Nagai, M. Asada, and M. Myowa-Yamakoshi, "Infant's action skill dynamically modulates parental action demonstration in the dyadic interaction," *Developmental science*, 2014.
- [77] D. Kraft, N. Pugeault, E. BAŞESKI, M. POPOVIĆ, D. KRAGIĆ, S. Kalkan, F. Wörgötter, and N. Krüger, "Birth of the object: Detection of objectness and extraction of object shape through object–action complexes," *International Journal of Humanoid Robotics*, vol. 5, no. 02, pp. 247–265, 2008.
- [78] J. Modayil and B. Kuipers, "The initial development of object knowledge by a learning robot," *Robotics and autonomous systems*, vol. 56, no. 11, pp. 879–890, 2008.
- [79] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.
- [80] L. Lukic, A. Billard, and J. Santos-Victor, "Modulating vision with motor plans: A biologically-inspired efficient allocation of visual resources," in *Proceedings of the IEEE-RAS International Conference on Humanoid Robots*, no. EPFL-CONF-188246, 2013.
- [81] G. Metta, G. Sandini, and J. Konczak, "A developmental approach to visually-guided reaching in artificial systems," *Neural networks*, vol. 12, no. 10, pp. 1413–1427, 1999.
- [82] J. W. Hart and B. Scassellati, "Robotic self-models inspired by human development," in *Metacognition for Robust Social Systems*, 2010.
- [83] F. Chao, X. Zhang, H.-X. Lin, C.-L. Zhou, and M. Jiang, "Learning robotic hand-eye coordination through a developmental constraint driven approach," *International Journal of Automation and Computing*, vol. 10, no. 5, pp. 414–424, 2013.
- [84] P.-Y. Oudeyer and F. Kaplan, "What is intrinsic motivation? a typology of computational approaches," *Frontiers in neurorobotics*, vol. 1, pp. 1–6, 2007.
- [85] E. Ugur and J. Piater, "Emergent structuring of interdependent affordance learning tasks," in *IEEE Intl. Conf. on Development and Learning and on Epigenetic Robotics (ICDL-Epirob)*, 2014.
- [86] J. Piater, S. Jodogne, R. Detry, D. Kraft, N. Krüger, O. Kroemer, and J. Peters, "Learning visual representations for perception-action systems," *The International Journal of Robotics Research*, vol. 30, no. 3, pp. 294–307, 2011.
- [87] N. A. Bernstein, "The co-ordination and regulation of movements," 1967.
- [88] M. Lungarella and L. Berthouze, "Adaptivity via alternate freeing and freezing of degrees of freedom," in *Neural Information Processing, 2002. ICONIP'02. Proceedings of the 9th International Conference on*, vol. 1. IEEE, 2002, pp. 482–487.
- [89] S. Szedmak, E. Ugur, and J. Piater, "Knowledge propagation and relation learning for predicting action effects," in *IEEE Intl. Conf. on Intelligent Robots and Systems (IROS)*, 2014.
- [90] M. Do, J. Schill, J. Ernesti, and T. Asfour, "Learn to wipe: A case study of structural bootstrapping from sensorimotor experience," in *Proceedings of IEEE International Conference on Robotics and Automation*, 2014.
- [91] E. E. Aksoy, M. Tamosiunaite, R. Vuga, A. Ude, C. Geib, M. Steedman, and F. Worgötter, "Structural bootstrapping at the sensorimotor level for the fast acquisition of action knowledge for cognitive robots," in *IEEE Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL)*, 2013, pp. 1–8.
- [92] E. Ugur, S. Szedmak, and J. Piater, "Bootstrapping paired-object affordance learning with learned single-affordance features," in *IEEE Intl. Conf. on Development and Learning and on Epigenetic Robotics (ICDL-Epirob)*, 2014.
- [93] S. J. Pan and Q. Yang, "A survey on transfer learning," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 22, no. 10, pp. 1345–1359, 2010.
- [94] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 2, pp. 265–286, 2007.
- [95] B. Moldovan, P. Moreno, M. van Otterlo, J. Santos-Victor, and L. De Raedt, "Learning relational affordance models for robots in multi-object manipulation tasks," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4373–4378.
- [96] S. Schaal, "Dynamic movement primitives—a framework for motor control in humans and humanoid robotics," in *Adaptive Motion of Animals and Machines*. Springer, 2006, pp. 261–280.
- [97] A. Ude, A. Gams, T. Asfour, and J. Morimoto, "Task-specific generalization of discrete and periodic dynamic movement primitives," *Robotics, IEEE Transactions on*, vol. 26, no. 5, pp. 800–815, 2010.
- [98] F. Stulp, E. A. Theodorou, and S. Schaal, "Reinforcement learning with sequences of motion primitives for robust manipulation," *Robotics, IEEE Transactions on*, vol. 28, no. 6, pp. 1360–1370, 2012.
- [99] E. Ugur and J. Piater, "Bottom-up learning of object categories, action effects and logical rules: From continuous manipulative exploration to symbolic planning," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015.