

Erhan Oztop · Nina S. Bradley · Michael A. Arbib

Infant grasp learning: a computational model

Received: 30 December 2002 / Accepted: 24 March 2004 / Published online: 25 June 2004
© Springer-Verlag 2004

Abstract This paper presents ILGM (the Infant Learning to Grasp Model), the first computational model of infant grasp learning that is constrained by the infant motor development literature. By *grasp learning* we mean learning how to make motor plans in response to sensory stimuli such that *open-loop* execution of the plan leads to a successful grasp. The open-loop assumption is justified by the behavioral evidence that early grasping is based on open-loop control rather than on-line visual feedback. Key elements of the infancy period, namely elementary motor schemas, the exploratory nature of infant motor interaction, and inherent motor variability are captured in the model. In particular we show, through computational modeling, how an existing behavior (reaching) yields a more complex behavior (grasping) through interactive goal-directed trial and error learning. Our study focuses on how the infant learns to generate grasps that match the affordances presented by objects in the environment. ILGM was designed to learn execution parameters for controlling the hand movement as well as for modulating the reach to provide a successful grasp matching the target object affordance. Moreover, ILGM produces testable predictions regarding infant motor learning processes and poses new questions to experimentalists.

Keywords Grasp learning model · Infant active exploration · Infant motor behavior · Motor development · Motor learning · Reinforcement learning

E. Oztop (✉)
ATR, Computational Neuroscience Laboratories,
2-2-2 Hikaridai, Seika-cho,
619-0288 Soraku-gun, Kyoto, Japan
e-mail: erhan@atr.jp

E. Oztop · M. A. Arbib
USC Brain Project, University of Southern California,
Los Angeles, CA 90089-2520, USA

N. S. Bradley
Department of Biokinesiology & Physical Therapy, University
of Southern California,
1540 E. Alcazar St., CHP155,
Los Angeles, CA 90033, USA

Introduction

Although it was recently suggested that human grasping is a generalized reaching movement encompassing the fingers so as to bring them to their targets on the object surface (Smeets and Brenner 1999, 2001), the general view is that separate but coordinated controllers execute mature hand transport and finger movements (see Jeannerod et al. 1998). Neither of these hypotheses explains the development of grasping behavior from infancy to adulthood. Although there are many experimental studies on the attributes of infant reaching and grasping and a number of models of reaching to grasp (Arbib and Hoff 1994), there are no computational models of infant grasp learning that combine the results of the empirical studies and pose new questions to experimentalists. Following Meltzoff and Moore (1977) it is assumed that infants are able to imitate facial and manual gestures soon after birth. Imitation of object manipulation has been reported for 14-month-old infants (Meltzoff 1988). However, it is yet unknown whether infants (as young as 4 months old) develop grasping ability via imitation or by self-regulated learning mechanisms. In this paper, we review evidence from literature that visual processing of hand-object relation is not developed at early grasping ages, and build ILGM, the Infant Learning to Grasp Model, based on the literature to answer the question whether grasping skill can be acquired with limited visual processing (that forbids a delicate object-related imitation capability.) Then we investigate the properties of ILGM grasp learning for validating the model and making predictions for the experimentalists. In particular, we analyze the effect of environment and task constraints in shaping infant grasping skill. The model is built on the notion that infants initially have a concept neither of minimizing an ‘error’ between hand shape and object shape, nor of how to match their hands to objects for grasping. The model shows that two assumptions suffice for an infant to learn effective grasps: 1) infants sense the effects of their motor acts, and 2) infants are able to use feedback to adjust their movement planning parameters.

The basis for motor learning is established early in development. Infants begin to explore their bodies, investigating intermodal redundancies, and temporal and spatial relationships between self-perceptions by 2–3 months of age (Rochat 1998). By 3–5 months, infants are aware of the visual, proprioceptive and haptic consequences of their limb movements (Rochat and Morgan 1995). Children learn the possibilities of action and the affordances for action in the environment through such exploratory behavior (Olmos et al. 2000). In particular, the affordances for object manipulation include the visual or tactile cues indicating that an object or a portion of it constitutes a suitable target for a stable grasp. Infants also learn to overcome problems associated with reaching and grasping by interactive trial and error learning (von Hofsten 1993; Berthier et al. 1996). They learn how to control their arms and to match the abilities of their limbs with affordances presented by the environment (Bernstein 1967; Gibson 1969, 1988; Thelen 2000). Our study focuses on the latter skill, how to learn to generate grasps that match the affordances presented by objects in the environment. By *learning to generate grasp* we mean learning how to make motor plans in response to sensory stimuli such that execution of the plan without visual feedback can lead to a successful grasp (without denying the importance of visual feedback in more elaborate manual tasks).

The basic motor elements for reach to grasp are established early in development, and early motor learning leads to rapid development of sensorimotor strategies during the first year. Infants exhibit a crude ability to reach at birth (von Hofsten 1982), which evolves into distinct skills for reaching and grasping by 4–5 months, and adult-like reaching and grasping strategies by 9 months (von Hofsten 1991), achieving precision grasping by 12–18 months (von Hofsten 1984).

Reaches of the neonate elicited by vision of a target can be executed without vision of the hand because infants 12 weeks of age make hand contact with glowing and sounding objects under lighted and dark conditions with similar frequency, and the onset of successful grasping under the two conditions occurs at approximately the same age of 15–16 weeks (Clifton et al. 1993). Furthermore, experiments controlling the visual information (i.e., hand vision and target vision) available to 9-month-old infants during grasping show that visual monitoring of hand and target is not necessary for online adjustment of the hand to match target object orientation (McCarty et al. 2001). Yet, infants are apparently also learning to use visual information for monitoring as well as preplanning reaches because by 5 months of age, unexpected visual blockade of the hand disrupts the reach (Lasky 1977). Around 5–7 months, hand closure begins near the time of object contact (von Hofsten and Ronnqvist 1988) and hand orientation is partially specified relative to the vertical versus horizontal orientation of an object prior to contact (von Hofsten and Fazel-Zandy 1984) or by the time of contact (McCarty et al. 2001). Between 9–13 months of age, reaches appear more distinctly pre-programmed, as evidenced by earlier

hand orientation and anticipatory grasp closure relative to object orientation and size (Lockman et al. 1984; von Hofsten and Ronnqvist 1988; Newell et al. 1993). While both power and precision grips are observed during reaches at 6 months of age, the precision grip gradually becomes the dominant form, appropriately modified for smaller (pincer grasp) versus larger objects (forefinger grasp) during the second year (Butterworth et al. 1997). Collectively, these studies suggest that infants use somatosensory information to increase the success of early reaching and grasping efforts and to establish a basis for visually elicited (but not so much guided) reaching to grasp during the later portion of the first postnatal year. Later in development, deployment of initial visually elicited grasping probably provides the training stimuli for development of visual-feedback-based grasping required for delicate manipulation and accommodation of unexpected perturbations of the target and/or trajectory.

In this study, we tested various hypotheses and made testable predictions on the process of motor learning in infancy using computational modeling. The Infant Learning to Grasp Model (ILGM) we developed embraces the learning-based view of motor development without denying the necessity of a mature neuromuscular substrate. Our model interacts with its environment (plans and executes grasp actions), observes the consequences of its actions (grasp feedback), and modifies its internal parameters (corresponding to neural connections) such that certain patterns (grasp plans) are selected and refined. Based on the available experimental literature, we propose that infant grasp learning is mediated by neural circuits specialized for grasp planning which can function with only limited visual analysis of the object, i.e., its position and a crude assessment of grasp affordance. In the macaque monkey, a specialized circuit in the parietal area AIP extracts object affordances relevant to grasping (Taira et al. 1990; Sakata et al. 1998, 1999; Murata et al. 1996, 2000) and relays this information to the premotor cortex where contextual and intention related bias signals are also integrated for grasp selection/execution (see Fagg and Arbib 1998). It is very likely that a similar circuit exists in the human (see Jeannerod et al. 1995) and is adapted during infancy for subsequent acquisition of adult grasp skills. Our model presents mechanisms for the functional emergence of grasp circuits with training.

With ILGM we have performed four groups of simulation experiments. Simulation Experiment 1 (SE1) tested whether ILGM could emulate the first stage of infant grasping skill by learning to grasp using goal-directed variable reaches. SE1a assumes an automatic palm orientating function, while SE1b removes this assumption to test whether ILGM can learn the appropriate palm orientations required for stable grasps. In Simulation Experiment 2 (SE2) we tested the hypothesis that context-enforced task constraints could be a factor responsible for the infants' transition from power grasp to precision grip during development by having ILGM interact with a small cube either placed on or under a horizontal plane. Simulation Experiment 3 (SE3) was designed to test the

controlled effect of affordance availability to ILGM and compare simulation findings to those for infant grasping which show that although infants younger than 9 months of age can grasp an object once they contact it, they appear not to use the object affordance information (e.g., orientation) in grasp planning and execution. Finally, Simulation Experiment 4 (SE4) tested execution generalization to different target locations via training of ILGM with random object locations in the workspace. In the remainder of the paper, we specify ILGM in detail and analyze the model's relevance via simulation experiments. We present the resulting behavioral responses and make comparisons where experimental human data are available. We discuss predictions generated by the simulations that can be experimentally tested.

Methods

All studies with human and animal subjects at the University of Southern California do require prior IRB approval. However, the experiments conducted in preparing this paper were all carried out as

computer simulations. Such studies do not require approval of the University's Internal Review Board. The data from human and animal subjects used to evaluate our simulation results are all taken from the literature.

The simulation environment created is composed of 1) a model of the infant hand and arm based on the kinematics model of Oztop and Arbib (2002), 2) the computer routines implementing the infants' elemental hand behaviors, and 3) the actual implementation of ILGM learning routines.

In general, a simulation session is set up as follows. ILGM generates a set of grasp parameters based on the probability distribution represented in its layers (detailed below). However, with a small probability a random plan can be generated. The randomness captures the infant's variable and spontaneous movements and provides the exploration required for learning. When testing, the model generates the grasp plans solely based on the probability distribution represented in the layers. The task of learning is then to adjust the weights determining the layer probability distributions so that successful grasp parameters are represented with higher probability. The model receives positive rewards from those grasps that yield successful or nearly successful grasps, and negative rewards for plans yielding unstable grasps or no object contact. Thus, an almost-grasp program is encouraged, leading to a greater chance of producing similar plans on subsequent trials.

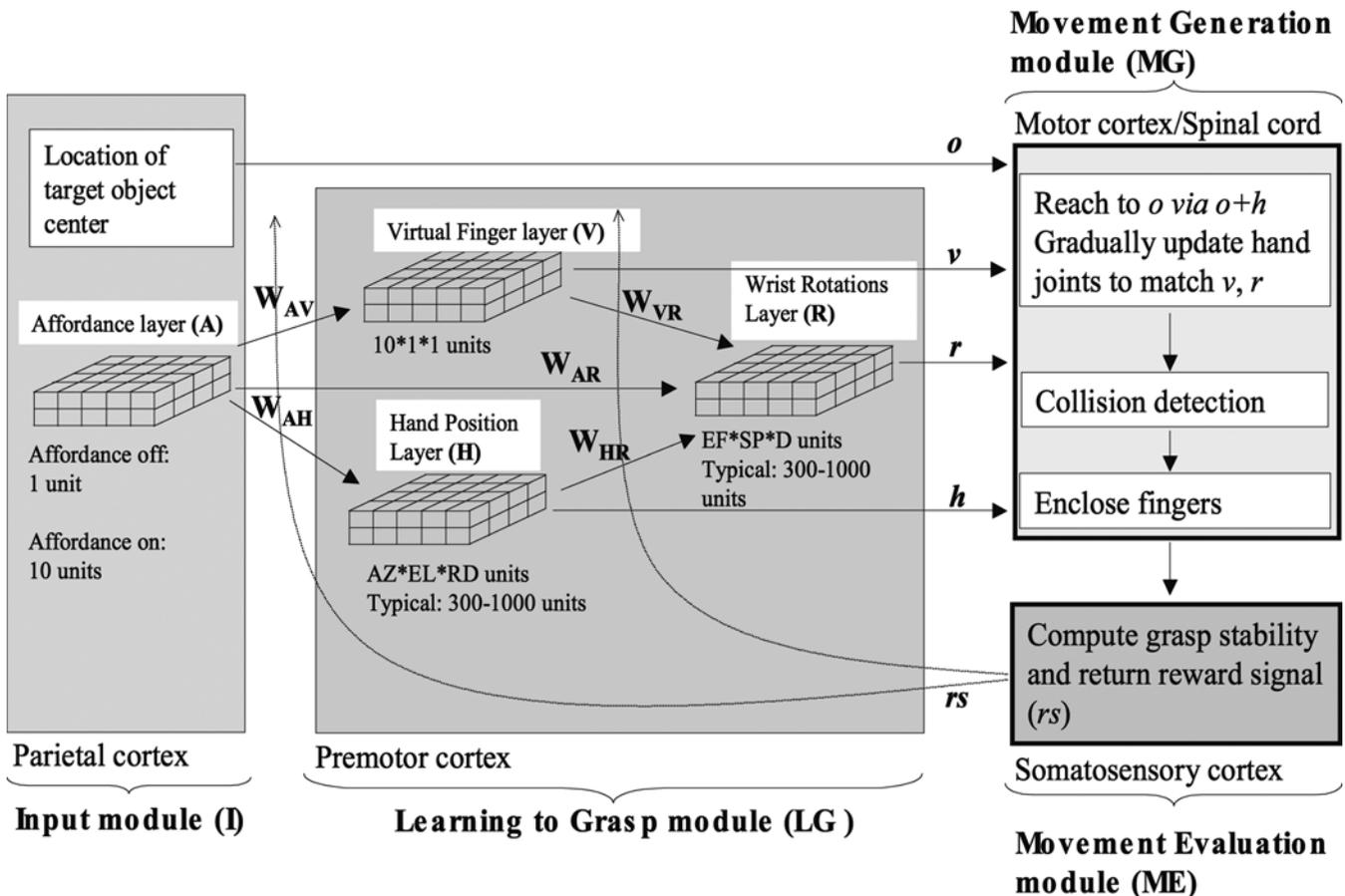


Fig. 1 The structure of the Infant Learning to Grasp Model. With o the location of the center of the target object, the aim of learning in the LG module is to provide inputs v , r and h to the MG module such that $o + h$ is a good via-point for a grasp based on virtual finger vector v and wrist rotation vector r to successfully grasp the object at o when it has the affordance encoded by affordance layer A. The individual layers are trained based on somatosensory feedback (AZ azimuth, EL elevation, RD radius; EF extension/flexion, SP

supination/pronation, D radial/ulnar deviation). The affordance layer is used to convey the affordances available to the model. In the simplest case, the model has only the information about the existence of an object (1 unit). In the most general case, it may include any object information, including the location. In the simulations reported in this paper, the affordance layer is engaged in encoding either the existence or the orientation or the position of the target object

Infant learning to grasp model (ILGM)

In our model, initiation of the reach is programmed, not learned, for our concern was to develop a model of grasp learning that shapes reaching movements into grasping movements. However, ILGM does learn to coordinate the reach with the grasp parameters it generates as a result of interactions with target object(s). These design elements are consistent with infants' real life experiences in that they begin grasping objects before reaching ability is fully developed, and each imperfect reach provides the opportunity to explore the space surrounding the object.

The schema level architecture of the Learning to Grasp Model is illustrated in Fig. 1. The left side panel shows the input (I) module providing sensory information. We assume that the general details of neural architecture for the reach and grasp are shared by human and monkey brains. Thus, while the behavior of the ILGM is designed to learn grasping in a manner consistent with data from human infants, the layers of the model are associated with general brain areas in a fashion based on monkey neurophysiology data. In brief, the input module (I), located in the parietal lobe, extracts object affordances and relays them to the learning to grasp module (LG), assumed to be in premotor cortex, the center implicated in grasp programming (Taira et al. 1990; Sakata et al. 1998, 1999; Murata et al. 1996, 2000). In turn, premotor cortex makes a grasp plan and instructs the movement generation module (MG) in the spinal cord and motor cortex for task execution (see Jeannerod et al. 1995). The sensory stimuli generated by the execution of the plan are integrated in the primary somatosensory cortex which contains the movement evaluation module (ME). Output of the somatosensory cortex mediates adaptation of the grasping circuit. The LG module's *Virtual Finger*, *Hand Position* and *Wrist Rotation* layers are based on the *Preshape*, *Approach Vector* and *Orient* grasping schemas proposed by Iberall and Arbib (1990). These layers encode a minimal set of kinematic parameters specifying basic grasp actions.

During early visually-elicited reaching while infants are reaching towards visual or auditory targets, they explore the space around the object and occasionally touch the object (Clifton et al. 1993). ILGM models the process of grasp learning starting from this stage. We represent infants' early reaches using an object-centered reference frame and include random disturbances during initial reach trials by ILGM so that the space around the virtual object is explored. We posit that the *Hand Position* layer specifies the hand approach direction relative to the object, realizing that the object can be grasped and touched from many angles. Intuitively speaking, the *Hand Position* layer contains the information for determining how to approach the object from, e.g., the top, bottom or right in a continuum of possible directions. A spherical coordinate system (instead of a Cartesian system) centered on the object center was chosen for the *Hand Position* layer, for its robustness in specifying an approach direction in the face of perturbations in vector components (see Fig. 14 in Appendix 1). More generally, the coordinate system may be centered on a particular affordance, rather than the object center.

Wrist orientation depends crucially on where the object is located relative to the shoulder as well as on how the reach is directed towards the object. Thus, the task of ILGM is to discover which orientations and approach directions are appropriate for a given object at a certain location, emphasizing that both an object's intrinsic properties and also its location create a context in which the grasp can be planned and executed. Thus, the *Wrist Rotation* layer learns the possible wrist orientations given the approach direction specified by the *Hand Position* layer. The *Wrist Rotation* layer also receives projections from the *Affordance* layer, because in general, different objects afford different paired sets of approach-direction and wrist-rotation solutions. The parameters generated by this layer determine the movements of hand extension-flexion, hand supination-pronation, ulnar and radial deviation.

The *Virtual Finger* layer indicates which finger synergies will be activated given an input. This layer's functionality is fully utilized in adult grasping and is engaged in only learning the synergistic enclosure rate of the hand in the current simulations. This layer can account for the developmental transition from synergistic finger

coordination during grasping (Lantz et al. 1996) to selective digit control for matching hand shape to object shape (Newell et al. 1993). In the simulations presented, the output of the *Virtual Finger* layer is a scalar value determining the rate of enclosure of the fingers during the transport phase. However, this does not restrain us from reproducing infant behavior and generating testable predictions.

Joy of grasping

Infants spontaneously engage in manual manipulation. Infants as young as 2 months of age play with their own hands, manipulate objects put in their hands, and play with rattles (Bayley 1936). When the hand contacts a glowing or sounding object in the dark, infants as young as 11 weeks try to grasp it (Clifton et al. 1993). We suggest that the resulting tactile stimuli yield neural signals that motivate infants to engage in grasping and holding. We further propose that the sensory feedback arising from the stable grasp of an object, 'joy of grasping', is a uniquely positive, motivating reward for the infant to explore and learn actions that lead to grasp-like experiences. This may be akin to the *adaptive value* of an action proposed by Sporns and Edelman (1993) when postulating three concurrent steps for sensorimotor learning: 1) spontaneous movement generation; 2) ability to sense the effects of movements and recognize their adaptive value; and 3) ability to select movements based on their adaptive value. Along these lines, we present a probabilistic neural network architecture that uses 'joy of grasping' as the reward stimulus and employs adaptive mechanisms similar to those used in reinforcement learning (see Sutton and Barto 1998 for a review). However, the goal of ILGM is not simply to find action which returns the maximal reward but rather to discover the repertoire of rewarding grasp actions in the form of probability distributions over the individual grasp parameters. This structure allows fast subsequent adaptations (e.g., inhibition of a hand position vector will block the generation of compatible wrist rotation vectors without retraining).

Movement parameter representation: learning to grasp (LG) module

This section describes the encoding used within the LG layers shown in Fig. 1.

Virtual finger (V) layer

In the simulations, the *Virtual Finger* layer generates a scalar value controlling the rate of flexion of the finger joints during movement execution. The MG module uses this value to implement the enclosure phase for the virtual hand during reach (see Movement generation). The minimal value, zero, informs the MG module that the hand will move without any enclosure, while the maximal value, one, indicates that the hand will move with an enclosure speed such that when the reach is complete the hand will be fistled. (Of course, this only happens if the reach does not produce manual contact with the target object.) The V layer is formed by a linear array of units each of which has a preferred enclosure value. The net movement parameter is calculated using a population-coding scheme: the activity-weighted sum of preferred values is returned (see Movement parameter generation and Learning) as the parameter represented.

Hand position (H) layer

In reaching for a target whose center is at position o , the hand must reach a via-point outside the target from which it moves toward the target. Let us represent that initial position by the vector sum $o + h$.

The H layer produces this offset vector, h . The parameter h is composed of azimuth, elevation and radius components. Each unit in the H layer thus has a preferred (azimuth, elevation, radius) value. The topology of the units forms a three-dimensional mesh. The net movement parameter is calculated using a population-coding scheme (see Movement parameter generation and Learning).

The MG module combines the vector o (available to it as a vector without neural encoding) and the offset h and then takes the hand to a via-point at $o+h$ and then to the object center o . (Note that this movement will often be interrupted by the collision of the hand with the object, triggering the enclosure reflex.) This simple scheme captures the basic ability to approach a goal from a range of directions.

Wrist rotation (R) layer

The R layer produces a vector, r , which informs the arm/hand simulator of the extent of the wrist movements to be made during the reach. The wrist movements indicated with r are hand extension/flexion, hand supination/pronation, ulnar/radial deviation. Similar to the H layer, each unit has preferred extension/flexion, hand supination/pronation and ulnar/radial deviation values. Thus, the topology of the units forms a three-dimensional mesh. The net movement parameter is calculated using a population-coding scheme as in the H layer. The reach simulator uses the r -value to rotate the hand via a linear interpolation between the current wrist angles and the r -value during the execution of the reach.

Movement generation: MG module

Since the goal of our study was to model how infant reaching movements can be shaped into grasping movements, reaching was not learned but implemented using techniques from robotics. More specifically, the Jacobian transpose method (see Sciavicco and Siciliano 2000) was used to solve the inverse kinematics (Appendix 2). As described in Appendix 2, our hand model has many degrees of freedom and thus its interaction with the target object can yield many contact situations. Figure 2 shows a simple contact scenario with two contact points. In general, each segment of the fingers can potentially create object contact. Our formulation takes into account the case of multiple contacts but uses the assumption that each one is a point contact. The list of contacts points (*contact list*) is provided to the movement evaluation module (ME) which computes the success value emulating the somatosensory feedback that we call the ‘joy of grasping’.

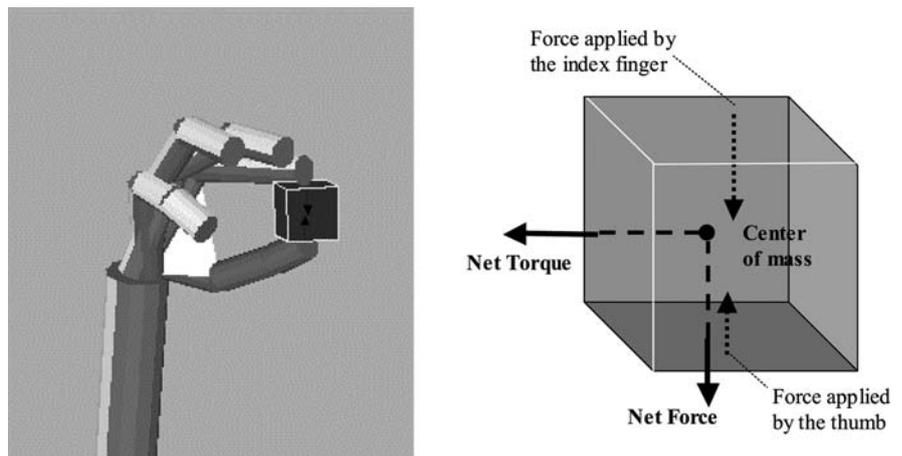
One of the simulation decisions is which part of the hand to use as the end-effector (see Appendices 2 and 3). Empirically we observed that reaches with (any part of) the thumb as the end-effector allow very limited grasp learning. The index and the middle finger are

most suited for ILGM (and indeed a study of neonates by Lantz et al. [1996] reports that the index finger is the first to contact an object). To keep the learning time required low, we did not implement another layer to learn end-effector assignment but rather specify it as a simulation parameter (Appendix 3). Figure 1 shows the inputs available to the MG module. Except for the o vector (object center), the parameters are generated by neural layers of the LG module. The algorithm for the execution of the movement is as follows:

1. Initialization: the vectors: o (object center), h (approach offset), r (hand rotations), θ_c (current arm joint configuration), r_c (current wrist joint configuration) and v (enclosure rate)
2. While the via point $o+h$ is not reached
 - a. Compute joint angle increment ($\Delta\theta_{o+h}$) to reach $o+h$ (Appendix 2)
 - b. $\theta_c \leftarrow \theta_c + \Delta\theta_{o+h}$
 - c. (Linearly) interpolate wrist joint angles $r_c \rightarrow r$
3. While o is not reached or enclosure is incomplete
 - a. Compute joint angle increment ($\Delta\theta_o$) to reach o (Appendix 2)
 - b. $\theta_c \leftarrow \theta_c + \Delta\theta_o$
 - c. Continue interpolating wrist joint angles $r_c \rightarrow r$
 - d. Flex finger joints with rate v
 - e. If target angles are achieved go to 4
 - f. Check for hand-object collision, if collision, execute enclosure reflex (Step 5)
4. Return empty contact points list. End
5. Independently flex each finger joint until object surface contacted or the joint limit is hit
6. Return the list of contact points
7. End

Appendix 2 presents the details of the simulated arm model together with the algorithm used to compute Steps 2a and 2b. In the algorithm above, the wrist joints are activated at the onset of reach (Step 2). However, for learning purposes this is not crucial; the algorithm works equally if the wrist joints were activated at a later time during hand transport. Remember that the aim of the algorithm is not to replicate human reaching (i.e., we are not concerned with for example generating bell shaped velocity profiles observed in adults) but rather to provide a wide range of hand object contact possibilities to mimic the infant’s variable reaches and grasp attempts, which mediate learning. ILGM’s learning routine will exploit whatever successful grasps occur to learn appropriate grasps to use when the hand approaches the object in the manner specified by MG.

Fig. 2 The grasp stability used in the simulations is illustrated for a hypothetical precision pinch grip in which the contact list has two elements, one for the thumb and one for the index finger



Mechanical grasp stability: movement evaluation (ME) module

The ME module takes the contact list returned by the MG module and returns a reward (or success) signal to be used for adapting the LG layer. A successful grasp requires that the object remains stable in the hand (it must not drop or move) (MacKenzie and Iberall 1994). Figure 2 illustrates a precision pinch performed by our simulator. To formalize this, let M be the number of contacts; P_i , N_i and F_i , respectively, the coordinates of contact i , the surface normal, and the force applied by the finger segment involved in the contact, and let P_c represent the coordinate of the center of mass of the object. Then grasp stability can be defined as follows (adapted from Fearing 1986) ($\langle \cdot \rangle$ and \otimes denote dot and cross products, respectively):

1. The net force (F_{net}) acting on the object must be zero where

$$F_{net} = \sum_{i=1}^M N_i \langle F_i \cdot N_i \rangle \quad (1)$$

2. The net torque (T_{net}) acting on the object must be zero where

$$T_{net} = \sum_{i=1}^M (P_i - P_c) \otimes N_i \langle F_i \cdot N_i \rangle \quad (2)$$

3. For any force acting on the object, the angle between the directions of the force and the surface normal must be less than a certain angle ρ . This angle is defined through the finger and contact surface dependent coefficient μ with the relation $\rho = \tan^{-1} \mu$ (i). The constant μ satisfies the property that if F_n and F_t are the normal and tangential components of an applied force then there will be no slip if $\mu |F_n| > |F_t|$ (ii). Combining (i) and (ii) we can write: there will be no slip if

$$\cos^{-1} \left(\frac{\langle F_i \cdot N_i \rangle}{|F_i| |N_i|} \right) < \rho \quad (3)$$

where ρ is a constant that is determined by the finger and contact surface dependent friction coefficient.

To evaluate how good (in terms of grasp stability) a hand-object contact is we may think of the ‘‘cost’’ of a contact list measured by a *cost function* indicating how much it violates the above three desiderata. However, since our simulator does not produce forces, we have to evaluate the contact list without force specification (i.e., have to evaluate the hand configuration). We perform this by checking whether a given contact list affords a set of contact-forces conforming to (1–3) by minimizing the cost function with respect to F_i . Then the grasp error is defined as the minimum value of the cost function. The following cost function (E) was used in the simulation experiments, which approximately captures the criteria (1–3).

$$E(F_i) = \alpha |F_{net}| + (1 - \alpha) |T_{net}| + \beta \sum_{i=1}^M \cos^{-1} \left(\frac{\langle F_i \cdot N_i \rangle}{|F_i| |N_i|} \right) \quad (4)$$

$$|F_i| > f_{min}, \forall i = 1, \dots, M$$

The constants α and β determine relative contributions of the force (α), torque $(1-\alpha)$ and non-slipping (β) cost terms to the total cost function. The first two terms capture the grasp stability conditions that the net force and net torque acting on the object must be zero (1, 2). The last term captures (approximately) the requirement of non-slipping contacts (3). The value f_{min} is an arbitrary small positive constant to avoid the degenerate solution $F_i=0$. Although the friction between the fingers and object surfaces is not explicitly specified, the non-slipping cost term ensures the rejection of grasp configurations in which the angle between contact

forces and the surface normals have high angular deviation. This means that the larger the β , the more slippery the object-finger contact, as evidenced by the fact that a large β will cause a large E for oblique forces acting on the object surface. Thus, while the α terms only include the vertical component of the acting forces, the tangential forces are included indirectly by the β term.

The grasp error (E_{min} , the minimum value of E with respect to finger forces) is converted into a reward signal (rs) to be used for adaptation in the LG layer as follows.

$$rs = \begin{cases} rs_{neg} & \text{if } E_{min} > E_{threshold} \\ 0.5 + e^{-E_{min}} & \text{if } E_{min} < E_{threshold} \end{cases} \quad (5)$$

where rs_{neg} is a negative, and $E_{threshold}$ is a positive simulation constant (see Appendix 3). The exact form of the reward signal generation is not important. What is sought is generation of a rs value that is negative for movements that are clearly not stable (large E_{min}), but positive for movements that are close to being stable ($E_{min} < E_{threshold}$), with increasing magnitude representing the stability of the grasp. Note that the special cases of no contact and single contact directly generate a negative reward. The constants are empirically determined in conjunction with the σ^2 parameter (see Movement parameter generation and learning) which controls the spread of activity to neighbor units during parameter generation and learning.

Movement parameter generation and learning: learning to grasp (LG) module

The LG layers are made up of stochastic units that represent probability distributions (see Zemel et al. 1998), but the parameter generated from a layer is determined by the unit that fires first (within a simulation step) as in rank order coding (van Rullen and Thorpe 2001). This probabilistic architecture enables representation of multiple action choices at the same time and allows biasing from other neural circuits. After learning, when presented with an object, the ILGM generates a menu that contains different grasp plans that are applicable to the affordances presented by the object. The menu could be biased by other brain regions such as the prefrontal cortex (Fagg and Arbib 1998) according to organism’s motivations or task constraints, but this is outside the scope of the present study.

The firing probability of a neuron is directly proportional to the sum of the weighted (by the connection matrix) input it receives (normalized over a layer). Our goal in the design of LG layers was to represent a *probability distribution* of grasp parameter values rather than directly representing the parameters. Hence, a grasp plan is determined by generating parameters based on the probability distributions represented in the layers. The goal of learning is to discover the probability distribution of grasp parameters that can achieve stable grasping. The grasp parameter generation and weight update algorithm given below computes the reward-weighted histogram of the output patterns as a function of input. In the simplest (1-layer) case, the action that returns reward causes a weight strengthening between the unit coding the input stimuli and the neuron coding the action. ILGM extends this 1-layer reward and input dependent histogram learning to more layers with the goal of eliminating parameter search around the incompatible parameters, thus speeding up learning.

The following description employs a number of notations from Fig. 1, and so the reader will find it helpful to refer to that figure while reading the following details.

Grasp parameter generation

In the following description, all the layers (A, V, H and R) are vectors (regardless of the encoded representation), the weights (W_{xy}) are matrices. The individual elements of

vectors are denoted by lowercase single letter subscripts (e.g., V_k).

Step 0

Encode the affordances presented to the circuit in Affordance layer (A) using population coding. Set the vector o as the center of the target object. In the simplest case the A is a single unit encoding the existence of an object (Simulation Experiments 1 and 2); in the most general case it encodes the features relevant for grasping such as orientation (Simulation Experiment 3) and position (Simulation Experiment 4).

Step 1

Compute the outputs of the Virtual Finger (V) and Hand Position (H) layers based on the connection matrices W_{AV} and W_{AH} (these connect I to LG; see Fig. 1):

$$V_{net} = [W_{AV}A]^+, H_{net} = [W_{AH}A]^+ \quad (6)$$

$\left\{ \begin{array}{l} \text{if (random} > \alpha) \\ \text{else} \end{array} \right.$	$\text{Pick } v^* \in [1, N_V],$ $\text{Pick } v^* \in [1, N_V],$	with the probability distribution represented by V with uniform probability distribution	
$\left\{ \begin{array}{l} \text{if (random} > \alpha) \\ \text{else} \end{array} \right.$	$\text{Pick } h^* \in [1, N_H],$ $\text{Pick } h^* \in [1, N_H],$	with the probability distribution represented by H with uniform probability distribution	(8)

Step 3

This step over-writes the outputs of the Virtual Finger (V) and Hand Position (H) layers from Step 2, now encode the “action values” h^* and v^* as a population code with variance σ^2 (a predetermined constant),

$$V_k = e^{-\frac{\delta(k,v^*)}{2\sigma^2}}, H_j = e^{-\frac{\delta(j,h^*)}{2\sigma^2}} \quad (9)$$

for $k=1 \dots N_V$ and for $j=1 \dots N_H$.

Here, $\delta(a,b)$ is the Euclidean distance metric defined over the indexes of a and b in the neural population mesh. In general, the function $\delta(a,b)$ measures how closely the units a and b contribute to a particular parameter value.

Step 4

Compute the net synaptic inputs into the Wrist Rotation layer (R) using the same conventions as in Step 1. N_R represents the number of units in R. (Again, see Fig. 1 for the placement of the connection weight matrices.)

where the $[\]^+$ operation rectifies the negative operands to zero; and

$$V = \frac{V_{net}}{\sum_{k=1}^{N_V} V_{net,k}}, H = \frac{H_{net}}{\sum_{k=1}^{N_H} H_{net,k}} \quad (7)$$

where N_V and N_H represent the number of units in layers V and H. This normalization step converts the net outputs into probabilities so that V and H can be treated as probability distributions.

Step 2

Process activity in layers V and H to extract v^* , h^* for next action. Below *random* represents a uniform random variable in $[0, 1]$. If *random* $< \alpha$ (where $0 < \alpha < 1$), a random output is generated, providing exploration in the parameter space. But if *random* $> \alpha$, then the probability distribution on V and H determine which output is chosen:

$$\begin{aligned} R_{net}^{synV} &= [W_{VR}V]^+, R_{net}^{synH} = [W_{HR}H]^+, \\ R_{net}^{synA} &= [W_{AR}A]^+ \end{aligned} \quad (10)$$

$$\begin{aligned} R^{synV} &= \frac{R_{net}^{synV}}{\sum_{k=1}^{N_R} R_{net,k}^{synV}}, R^{synH} = \frac{R_{net}^{synH}}{\sum_{k=1}^{N_R} R_{net,k}^{synH}}, \\ R^{synA} &= \frac{R_{net}^{synA}}{\sum_{k=1}^{N_R} R_{net,k}^{synA}} \end{aligned} \quad (11)$$

Step 5

Combine the inputs multiplicatively, then normalize them. (The intuition behind multiplication is that the probability distributions represented in V, H are a multiplicative factor of the probability distribution represented in R since the wrist rotation is a function of virtual finger, hand position and affordance input.)

$$R = \frac{R^{synV} * R^{synH} * R^{synA}}{\sum_{k=1}^{N_R} R_k^{synV} * R_k^{synH} * R_k^{synA}} \quad (12)$$

where the $*$ operator denotes component-wise multiplication.

Steps 6 and 7 now do for R what Steps 2 and 3 did for V and H.

Step 6

Obtain r^* from layer R.

$$\begin{cases} \text{if (random} > \alpha) & \text{Pick } r^* \in [1, N_R], & \text{with the probability distribution represented by R} \\ \text{else} & \text{Pick } r^* \in [1, N_R], & \text{with uniform probability distribution} \end{cases} \quad (13)$$

Step 7

Encode r^*

$$R_i = e^{\frac{-\delta(i,r^*)}{2\sigma^2}} \quad (14)$$

for $i=1 \dots N_R$.

Step 8

Extract the “population vector” to provide the actual values that drive the generation of the movement.

$$v = \sum_{k=1}^{N_V} P_V^k V_k, \quad h = \sum_{k=1}^{N_H} P_H^k H_k, \quad r = \sum_{k=1}^{N_R} P_R^k R_k \quad (15)$$

where P_X^k denotes the preferred movement parameter associated with unit k in layer X.

Step 9

Generate the reach using the parameters v , h and r and o (see Movement generation).

Step 10

Compute stability and generate reward signal rs (see Mechanical grasp stability).

Step 11: learning—adapt the weights using rs

The connection among layer units are updated with the REINFORCE update rule¹ (Williams 1992; Sutton and Barto 1998), which is Hebbian for positive rs and anti-Hebbian for negative rs values.

¹Technically, REINFORCE requires the firing probability of neurons be specified as a differentiable function of the input. Our algorithm does not conform to this criterion (i.e., Step 1) but it can be shown that a softmax approximation to Step 1 yields a learning rule similar to the one we used.

$$\begin{aligned} W_{AV}^{new} &= W_{AV} + \eta_{AV} rs AV^T, & W_{AH}^{new} &= W_{AH} + \eta_{AH} rs AH^T \\ W_{HR}^{new} &= W_{HR} + \eta_{HR} rs HR^T, & W_{VR}^{new} &= W_{VR} + \eta_{VR} rs VR^T \end{aligned} \quad (16)$$

The η denote the learning rate parameters, and super-script T denotes the vector transpose. Parameter values for our simulations are set out in Appendix 3.

Note that firing rates play a different role within LG from that which they play as outputs fed into the Movement Generation module (MG). Acting within LG, the initially determined firing rates across the V, H and R layers are used to then pick a specific unit in each layer, which will control the ensuing grasp (Steps 2 and 5). However, in preparing the outputs from LG that will affect MG, Step 3 over-writes the outputs of the Virtual Finger (V) and Hand Position (H) layers, recoding the “action values” from Step 2 as a population code for the chosen action value. Similarly, Step 7 over-writes the output of the Wrist Rotation layer (R), recoding the “action values” from Step 6 as a population code. Step 8 then takes these V, H and R population codes and extracts the “population vector” to provide the actual values v , h , and r that drive the generation of the movement. For a biological system, the over-writing mechanism implemented may represent a lateral inhibition mechanism within a layer, by which the first firing neuron determines the information to be transferred to subsequent layers (as suggested for visual cortex by van Rullen and Thorpe 2001).

Since the output is represented as a population code after Steps 3 and 7, respectively, the weight strengthening (Step 11) not only increases the probability that the neuron that coded the rewarding output will become active in future but also increases the chance of neighboring output units becoming active as well. Because the reward signals are used to update not only the units contributing to the action selection, but also their neighbors as well, large negative reward signals (rs_{neg}) may overpower the activity of a unit that yields positive reward. This happens when the parameters returning positive reward constitute only a small fraction of the whole parameter space. Therefore, the negative signal is usually chosen smaller than the positive reward signals (e.g., less than 0.1 in absolute value as compared with 0.5).

In summary, the Wrist Rotation layer receives the output of Hand Position and Virtual Finger layers and dynamically computes a probability distribution for wrist movement parameters based on the input signals. The dependency of wrist movement parameters on the input is captured in the connection matrices via weight adaptation

rule (Step 11). After learning, ILGM generates a set of grasp parameters as follows (see Fig. 1). Presented with the object, Hand Position layer produces the distribution of possible approach directions. The distribution is used to generate the hand approach direction vector that is relayed to Wrist Rotation layer. Similarly, the enclosure rate parameter generated by the Virtual Finger layer is relayed to Wrist Rotation layer. The Wrist Rotation layer in turn, computes the distribution of feasible wrist orientations given the specified approach direction, enclosure speed and the object affordance (output of Affordance layer). The probability distribution computed in the Wrist Rotation layer is then used to generate wrist rotation parameters completing the parameters of a grasp plan.

Simulation experiments

To test the hypothesis that early grasp learning could be mediated by simple goal-directed reaching without affordance extraction and imitation capability, we mimicked infant elemental hand behaviors, implementing reaching, grasp reflex and palm orienting behavior (programmed for Simulation 1, learned for other simulations). Each simulation experiment set was run independently with different settings, however linked by the conceptual progression of the hypotheses tested. For example, in Simulation Experiment 2 (SE2), we introduced an obstacle to create a ‘context,’ whereas in Simulation Experiment 4 (SE4) we varied object location to test the generalization of ILGM to different object locations. Consequently, the learning results of a simulation do not transfer to subsequent simulation experiments. However, the consistent conclusions drawn from our simulation results are presented in the Discussion. The period our simulations address (except SE3b), corresponds to the early postnatal grasping period when object affordances are not fully appreciated because of either maturational shortcomings (maturational theory) or developmental delays attributable to the complexity of learning (learning-based theories of motor development).

Simulation experiments 1 (SE1): learning to grasp with/without auto palm orienting

SE1a Although fisting appears to dominate open-hand reaches in the second month of age, the hands open again around the third postnatal month (von Hofsten 1984). During reaches with an open hand, infants orient their palms towards the target objects (see Streri 1993, pp 46–47). It has been suggested that the grasp reflex constitutes the first stage of grasping (Twitchell 1970; see Streri 1993 for a review). We designed our first simulation to verify that our model emulates the first stage of grasping skill by bootstrapping voluntary grasping based on goal-directed variable reaches and grasp reflex.

Given o , the location of the center of the target object, and the affordance encoded by affordance layer A, the

grasp plan provides a scalar v (*hand enclosure rate*), and two vectors, r (*3DOF rotations of the wrist*) and h (*3-dimensional offset vector*) to the MG module. The plan is such that $o+h$ is a good via point compatible with v and r to successfully grasp the object at o . However, in this simulation experiment (SE1a only), a grasp plan is defined by the single vector h which specifies the via-point $o+h$ that the hand will pass through (see Movement generation), since we program an automatic function that drives the fully extended hand with the palm always facing the target (see Appendix 2). The location is specified using a spherical coordinate reference frame centered on the (modeled) object (see Appendix 1). The hand goes first to the via-point, and then approaches the object. A contact with the object triggers the grasp reflex, which is modeled as a flexion of fingers over the object when the hand-object contact occurs. We view the grasp reflex as a spinal reflex, while viewing LG as modeling a crucial part of the cortical control of the grasping. In general, the grasp reflex need not yield stable grasping, as there is no guarantee that the model (or an infant) approached the object from an appropriate direction. Furthermore, although the object is suspended in space, the position of the object does not always allow grasping from all directions due to anatomical constraints. Obviously, the total range is not biologically feasible—but the model will learn which via points achievable by the simulated arm will support a stable grasp. An enclosure that yields a successful or nearly successful grasp generates a positive reward (see Mechanical grasp stability), whereas unstable grasps and lack of object contact yield negative reward.

After learning, the Hand Position Layer represents a probability distribution for the hand position vectors that lead to a suitable contact for a stable grasp. Since the Hand Position layer units encode azimuth, elevation and radius parameters, (α, β, r) , their activity creates a three-dimensional probability distribution. The quantization of the distribution is the result of the number of units allocated for each parameter. For this simulation, we used 726 units (a grid of $11 \times 11 \times 7$) (Appendix 3). By summing over the radius component, we project a grasp plan to a two-dimensional space and hence plot the probability as shown in Results, SE1a.

SE1b Although there are accounts that infants usually orient their hands during reaching to increase the likeliness of a contact between palm and the object (von Hofsten 1984), there is, to our knowledge, no evidence that this behavior is innate. The first aim of SE1b was to test whether the model could learn to orient the hand to produce appropriate object-contact for stable grasping (1). Thus in SE1b, we no longer use the automatic palm orienting capability, but instead test whether ILGM can learn the appropriate palm orientations required for stable grasps. The second aim of SE1b was to test whether ILGM could replicate the findings that in the early postnatal grasping period infants could initiate precision grips (2) (Newell et al. 1989; Butterworth et al. 1997). In terms of learning, the task of ILGM in this simulation was to

discover the distribution of all possible wrist movements (supination-pronation, extension-flexion, ulnar/radial deviation) that yield appropriate object-contact as a function of finger flexion rate and hand-object approach direction. The confirmation of (1) and (2) supports two hypotheses: (H1) during the early postnatal period infants can learn to orient their hand towards a target, rather than innately possessing it; (H2) infants are able to acquire grasping skills before they develop an elaborate adult-like object visual analysis capability, indicating that learning-by-imitation is not a prerequisite for grasp learning.

Simulation experiments 2 (SE2): task constraints shape infant grasping

If a variable (e.g., onset of precision grasping) is dynamically context specific, common experimental designs may be too artificial to reveal the dynamical nature of motor control (Bradley 2000). In Simulation Experiments SE2a and SE2b, we controlled the environment to observe the context specific nature of ILGM learning and tested the hypothesis that task constraints due to environmental context (or the action opportunities afforded by the environment) are a factor shaping infant grasp development (H3). If this hypothesis is correct, the observed general progression in infant prehension from power to dexterous precision grip (Halverson 1931; see Newell et al. 1989 for a review) can be explained as a process of reconciliation between increasing action possibilities (e.g., via postural development) with constraints introduced by novel task requirements and/or changing environment (e.g., a hanging toy versus a toy resting on a table). Task constraints may be viewed as including the goal of the task or the rules that constrain the response dynamics (Newell 1986). Some examples of task constraints are the object properties such as size and shape (Newell et al. 1989).

Infants of 5 months fail to retrieve an object when it is placed on a larger object, although they can retrieve an object when it stands out from the base (Diamond and Lee 2000). An infant's interaction with an object placed on a flat surface such as the floor or table is constrained by the base of the target as reach-to-grasp movements of the infant are interrupted by accidental collision with the target base (Diamond and Lee 2000). Thus, we tested hypothesis H3 by having ILGM interact with a small cube placed on a horizontal plane (SE2a: the *cube on the table* task.) In addition, to observe the differential effect of different task constraints, we repeated the experiment when the small cube was placed just under the horizontal plane (SE2b: the *cube under the table* task.)

Simulation experiments 3 (SE3): dowel grasping—the affordance-on/off cases

This simulation series was designed to test the effect of affordance availability to ILGM and compare simulation findings to those for infant grasping in a study by

Lockman et al. (1984), which presented infants 5 and 9 months of age with dowels in either the horizontal or vertical position. A successful grasp of an object with variable affordance (e.g., orientation) requires affordances to be incorporated into the grasp planning process. Although infants younger than 7 months of age can grasp an object once they contact it, they appear not to use the object affordance information. Hand orientation adjustments anticipate target rod orientation by mid-reach between 7–9 months (McCarty et al. 2001). Thus, we hypothesized that the inability of younger infants tested by Lockman et al. (1984) to pre-orient their hands could be explained by the lack of visual processing for extracting object affordances, rather than by motor immaturity. To test whether with this hypothesis ILGM yields results comparable with Lockman et al. (1984), ILGM was presented with a cylinder analogous to a dowel. We also included a third, diagonal orientation condition in addition to the existing horizontal and vertical orientation conditions. According to our hypothesis, we disabled the orientation encoding when simulating the younger infants' grasp learning (SE3a, *affordance-off* condition), and enabled the orientation encoding in the Affordance layer (as a population code) when simulating older infants' learning (SE3b, *affordance-on* condition). If our initial hypothesis is supported by the simulation results, then we can assert a stronger hypothesis: (H4) grasping performance in the 'affordance-off stage' mediates the development of a visual affordance extraction circuit. This extraction circuit is probably the homologue of monkey AIP (the anterior intraparietal area), an area known to be involved in computing object features relevant for grasping (Murata et al. 1996, 2000; Sakata et al. 1998, 1999). By detecting the common features of successfully grasped objects, infants may shape their visual analysis to emphasize the relevant features and represent those as compact neural codes (i.e., object affordances).

The Affordance (A) layer was modeled with 10 units encoding the orientation of the cylinder. In SE3b, the orientation of the cylinder was encoded in A; however, in SE3a each unit in A was set to 1. Since object affordances relevant for grasping were not available to ILGM in SE1, SE2 and SE3a, we associate these simulations with the infancy period of 2–6 months of age, and SE3b with the infancy period of 9–12 months of age.

Lockman et al. (1984) used orientation difference between the hand and the dowel as a measure of how much infants adapted their hand orientations to the target. If there was no orientation difference between hand and dowel, the trial was scored a 0, and maximal mismatch was scored a 4. Also, wrapping a finger around the dowel was counted as a grasp. When ILGM was learning in both Simulation 3 and Simulation 4, we used our grasp stability measure (see Mechanical grasp stability), but to compare our results with those of Lockman et al. (1984), we included unstable grasps that conformed to their grasp criteria. For example, the grasp in Fig. 7b does not satisfy

our grasp stability criterion but it does conform to criteria applied in Lockman et al. (1984).

Simulation experiment 4 (SE4): generalization to different target locations

In earlier simulation experiments, we were not concerned with the effect of object location and thus presented the objects always in the same location within a simulation session. Here we aim to show that when LG module has access to object location information, it can generate grasps as a function of object location. In this sense, the location of an object is part of the affordance of the object since the location (with other intrinsic object features) determines the grasps that are afforded. Thus, the object location is encoded in the Affordance layer as a population code and relayed to LG (for grasp learning/planning). The location was encoded using 100 (10×10) units in the Affordance (A) layer. The location was specified as a point on a (shoulder centered) sphere patch. We trained ILGM by randomly encoding the place of the target dodecahedron in the Affordance layer, in effect spanning the workspace as shown by the grid nodes in Fig. 13a. The generalization of the learning is tested by placing the target object furthest from the grid nodes used in training (i.e., the centers of the squares formed by the four equidistant nodes in the figure). Further we analyzed the dependency of the Hand Position layer output on the object location by plotting the Hand Position layer for a set of object locations [G1].

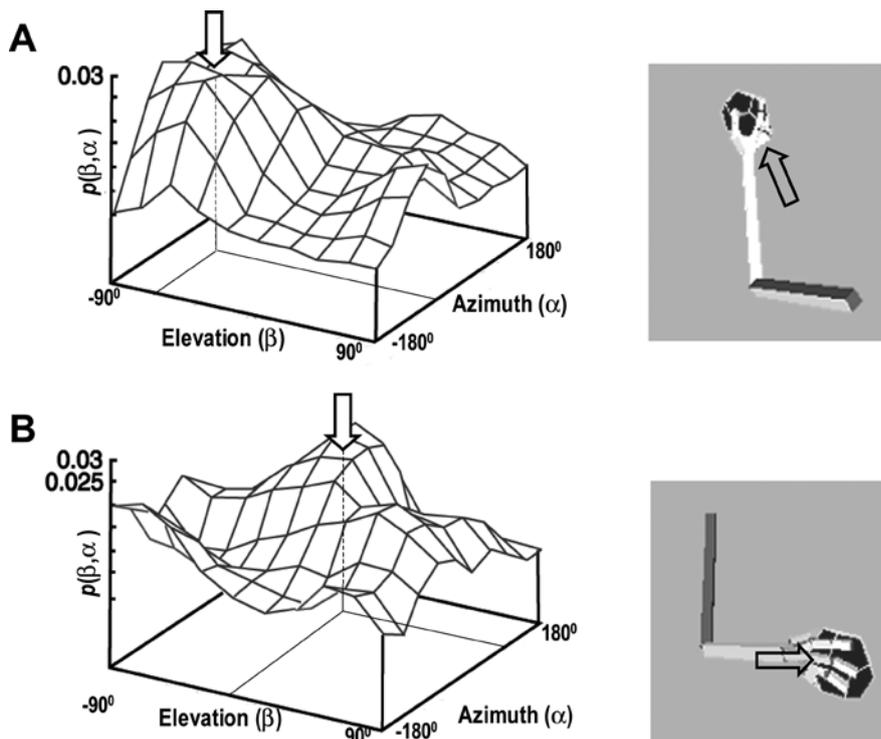
Results

SE1a: learning to grasp with auto-palm-orienting

We ran two simulations in which we positioned a sphere (modeled as a dodecahedron) in different locations in the workspace to see the differential learning for fixed target locations (but see SE4 for location dependent learning). Figure 3a illustrates a grasp discovered by the model for the case when the object was elevated in the workspace and the hand could not reach the space above the object. ILGM excluded grasping from the top, having discovered the arm could only grasp the elevated object from underneath and front/right. The azimuth, elevation and radius parameters (α , β , r) of the Hand Position layer create a three-dimensional probability distribution. We projected the distribution to the (α , β) space for the purpose of visualization (Fig. 3a) by summing over the radius component. This plot approximates the normalized histogram of (α, β) vectors generated by the model over many trials.

Note that, the elevation (β) range is between -90° and 90° , whereas the azimuth (α) range is between -180° and 180° as illustrated in Appendix 1. The plot shows that the model represented the approach directions $\beta < 0^\circ$ (below the object) and $-90^\circ < \alpha < 90^\circ$ (right and/or front of the object) with a higher probability, indicating that a grasp plan from this region will result in a stable grasp. The correspondence between the region with the highest peak and the approach direction is illustrated with *arrows* in Fig. 3a. We ran another simulation session after moving the object to a left lower area in the workspace of the modeled (right-handed) arm. In this case, we observed that

Fig. 3a, b (SE1a) The trained model's Hand Position layers are shown as a 3D plot for two different object positions. One dimension is summed to reduce the 4D map to a 3D map. The map shows that when the object is elevated in the frontal workspace a grasp *from bottom* is appropriate (a), whereas when the object is located on the left side of the workspace, an approach from the *right side* of the object would be appropriate (b). The correspondence between the most dominantly represented parameters and the corresponding approach directions are illustrated with *arrows* (a, b). The polar coordinate system in which the Hand Position layer parameters are specified is illustrated in Appendix 1



the model learned that a grasp directed to the right side of the object is likely to result in a stable grasp evidenced by the high activity where $\alpha > 0^\circ$. Figure 3b shows the results of this simulation.

SE1b: learning to grasp without auto-palm-orienting

Without auto-palm-orienting, the model learned to generate various grasp parameters to perform power and precision grasps, and many variations of the two. The most frequent grip generated was the power grip and its variations. Precision type grips were only occasionally generated.

ILGM simulation results are in accordance with empirical findings, when tested under proper conditions, infants are able to perform precision grips. The fact that the power grasp is inherently easier is evident in both infant performance and ILGM simulations: if the palm contacts the object, the sensory event is likely to produce a power grasp (i.e., palmer reflex). Figure 4a shows power grasp examples for a cube, a sphere (dodecahedron) and a cylinder. ILGM generated grasp plans for all three objects after separate learning sessions for each object.

The learned precision grasp varieties mainly engaged the thumb, index finger and one or two additional fingers; examples are illustrated in Fig. 4b. Usually the object was secured between the thumb and two or three fingers, the thumb opposing the fingers. The latter opposition (thumb opposing the index and middle finger) is in accordance with the theory of virtual fingers and opposition spaces (Iberall and Arbib 1990) and human tripod grasping (Baud-Bovy and Soechting 2001). Figure 4c shows examples of two finger precision grip that are less

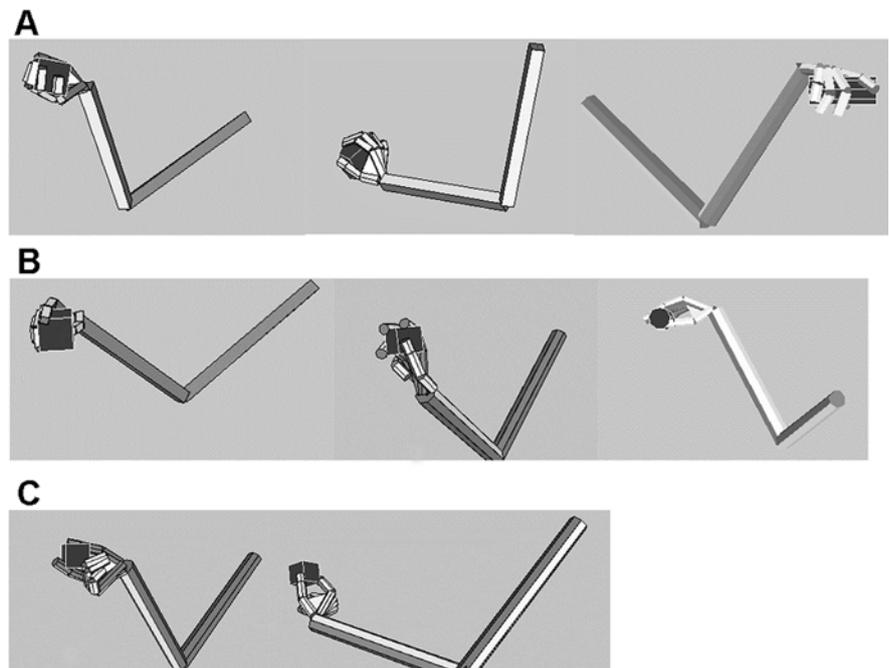
frequently generated compared with the three- or four-fingered precision grips.

Note the different wrist rotations that ILGM produced in response to the location of the hand with respect to the target object. ILGM learned to generate compatible hand location and wrist rotations to secure the object. Simulation results indicate that even without object affordance analysis, precision type grips can emerge from open-loop grasp learning, replicating findings on early infant grasping skills (Newell et al. 1989; Butterworth et al. 1997), and that the palm orienting behavior can emerge as a natural consequence of learning rather than being innately available to infants.

SE2a: cube on the table task

With the cube on the table task, the grasp repertoire formed by ILGM did not include power grasping (where the object is secured between palm and the fingers). Apparently power grasping was not represented in the repertoire because the fingers always collided with the modeled table surface and did not produce ‘joy of grasping’ signals (positive reward). The grasp ‘menu’ acquired for the cube on the table task was composed of grasps with wrist positions above the object. The contact points on the cube showed variability (see Fig. 5). Many of the precision grips that were learned enlisted supportive fingers in addition to the thumb and the index finger (Fig. 5a). However, two-finger precision grips were also acquired (Fig. 5b). Most of the precision grips learned correspond to inferior forefinger grasping (Fig. 5a) and inferior pincer grasping (Fig. 5b), based on the classification employed by Butterworth et al. (1997). One interesting observation is that ILGM represented the

Fig. 4a–c (SE1b) **a** After learning, ILGM planned and performed power grasps using different objects. Compare the supination (and to a lesser extent extension) of the wrist required to grasp the object from the bottom side (*left*) with the front (*center*) and upper side (*right*) grasp. **b** Two learned precision grips with a cube shaped object (*left* three fingered; *center* four fingered), and a precision grip learned for a cylinder (*right*) are shown. Note the different wrist rotations for each case. ILGM learned to combine the hand location with the correct wrist rotations to secure the object. **c** ILGM was able to occasionally generate two-fingered precision grips after learning; two examples are shown



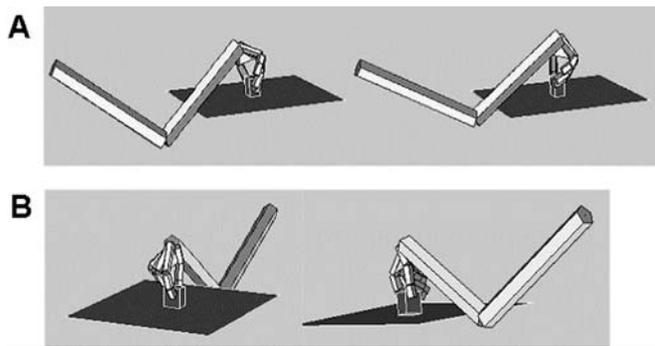


Fig. 5a, b (SE2a) ILGM learned a ‘menu’ of precision grips with the common property that the wrist was placed well above the object. The orientation of the hand and the contact points on the object showed some variability. Two example precision grips are shown in the figure. In addition to precision grips, where thumb opposes more than one finger (a), ILGM also acquired precision grips where the thumb opposes the index finger (b)

object affordances in the grasp ‘menu’ it learned. By comparing Fig. 5a and b, we see that the opposition axes used for grasping are 90° apart. In both forefinger grasps of Fig. 5a, the thumb was placed on the left surface of the cube, whereas during pincer grasping in Fig. 5b it is placed on the surface that is parallel and closer to the simulated shoulder.

SE2b: cube under the table task

The grasp repertoire formed by ILGM in the cube under the table task did not include power grasping, as in SE2a. ILGM learned precision type underarm grasps to avoid collision with the modeled table (see Fig. 6). One difference between cube on the table task and the cube under the table case was that the latter afforded a lateral pinch (Fig. 6b). This is apparently due to relative position of thumb and index finger with respect to the modeled table. With an underarm reach the thumb and the index finger could be brought as close as possible to the table surface, whereas with a reach from top, the thumb and the index finger cannot achieve arbitrary proximity to the surface (without an forceful awkward posture) disallowing grips for short objects such as the one we used in this simulation. By comparing the simulation results of SE2a and SE2b (cube on/under the table), we predict that underarm grasping is acquired late in development since cube under the table type of contexts are rare in infants’ daily life. Consequently, lateral pinch development could also be delayed compared to precision grip development. Each of these grasps also requires varying amounts of supination and Lockman’s study indicates forearm posture is biased into pronation so explorations likely have a pronation bias as well.

SE3a: dowel grasping—the affordance-off case

Figure 7 shows the orientations and the cylinder we used in the simulation. The grasp actions shown were performed by ILGM during the affordance-off condition. In this simulation, different cylinder orientations afford different hand orientations. Thus, we predicted that without affordance input, the model would not be able to learn a perfect grasping strategy. Using our grasp stability measure, observations indicated the horizontal cylinder was successfully grasped in six of ten trials. The vertical cylinder was grasped successfully in four out of ten trials, and the diagonal cylinder could not be grasped at all. Using the grasp criteria used by Lockman et al. (1984), the successful trials improved to: 8/10, 5/10, and 7/10, respectively.

Analyses of individual errors made by the model revealed six typical error curves (extent of mismatch) as shown in (Fig. 8). Note that curve *A1* is approximately constant and indicates there is no mismatch between the orientation of the hand and that of the cylinder, for both are horizontally oriented (0°), whereas in the other two curves of the same row (*B1* and *C1*), although the error is approximately constant, there is a 45° or 90° mismatch. The lower row plots have non-constant error curves (*A2*, *B2* and *C2*). The columns, from left to right, correspond to orientation mismatch of the hand with the dowel while grasping the horizontal cylinder, diagonal cylinder and vertical cylinder.

Lockman et al. (1984) reported that infants start their reaches with the hand horizontally oriented; thus, in our simulations the hand was set to a horizontal posture at reach onset causing a maximal mismatch for vertical cylinder grasping at the onset of movement. Consequently, matching the Lockman et al. (1984) observation, ILGM made more corrections during successful reaches for the vertical cylinder (Fig. 8, *C2*) and fewer for the horizontal cylinder (Fig. 8, *A1*). In the upper row of trials in Fig. 8 (*A1*, *B1*, *C1*), ILGM used a grasp plan appropriate for horizontal orientation as evidenced by the level of constant mismatch (0° for *A1*, 45° for *B1* and 90° for *C1*). As can be seen from the *C2* panel, the model can also occasionally perform a vertical cylinder adaptation (remember that the architecture of ILGM allows representation of multiple grasp plans). However, the model cannot differentiate the two grasping strategies. The model learns a strategy (probability distribution) that reflects the cylinder orientation presentation frequency. In the simulation, since the dowel orientation was selected randomly, ILGM represented both strategies (grasping a vertical and a horizontal dowel). However, the horizontal grasping strategy was favored as we observed that most of the errors made during grasp execution would fall into the constant error curve class (i.e., one of *A1*, *B1* or *C1*).

Fig. 6a, b (SE2b): underarm grasping learned. Most of the learned grasps were precision grips (**a**). Side or lateral grips (thumb opposing the side of the index finger) were also observed, as in **b**

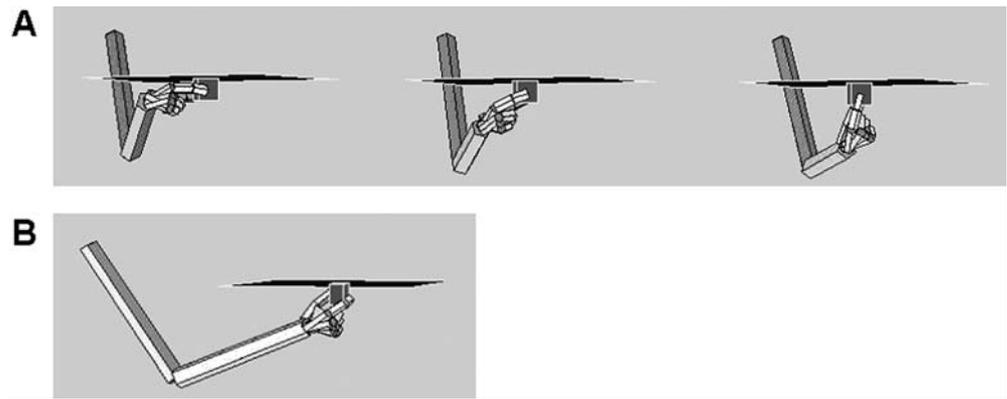
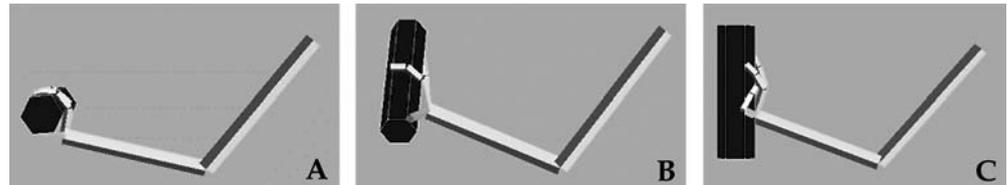


Fig. 7a–c (SE3a) the three cylinder orientations and grasp attempts of ILGM in the affordance-off condition. **b** and **c** Grasps do not satisfy our grasp stability criterion



SE3b: dowel grasping—the affordance-on case

First, ILGM was trained in the affordance-on condition. The grasping data after learning was then used with the grasping data from the affordance-off case learning (SE3a) to make a comparison with the data for 5- and 9-month-old infants in Lockman et al. (1984). For each condition, ILGM performed 64 grasping trials during presentation of the vertically oriented cylinder. We excluded the horizontal case (but see the Discussion later in this section), because for the horizontal cylinder ILGM learned a power grasp with hand supination (underarm grasp), a forearm posture not observed during reaches of infants tested by Lockman et al. (1984). Figure 9 shows the averaged orientation mismatch score for infants (a) and ILGM (b) at three locations as the hand approaches the cylinder. For capturing the manual scoring used by Lockman et al. (1984) we used the function $\text{round}(8^*x/\pi)$ to convert orientation differences (x) into mismatch scores. In the remainder of the figures, we use actual orientation difference (degrees) instead of mismatch score to avoid loss of information caused by scoring. The *dashed curve* in Fig. 9a depicts the performance of infants at 5 and 9 months. At 9 months of age (a, *solid curve*), infants more closely matched their hands to the vertical cylinder compared with infants at 5 months of age (Lockman et al. 1984). McCarty et al. (2001) controlled for vision of the hand and object and found the orientation was preset even if the light was extinguished after the onset of reach—suggesting that infants of 7–9 months needed only a brief view of the object to preplan the grasp.

Although the absolute scores differed between ILGM simulation and infant data, ILGM's affordance-on performance was better than the affordance-off case, analogous to differences in performance of infants at 9 versus 5 months of age. Moreover, the performance increments for ILGM and the infants were qualitatively comparable. The difference between human and simulation data at

movement onset is due to specification of starting hand orientation in the simulation in contrast to the variable orientation of infant hands (note that the scales of the plots in Fig. 9 are not same: the simulation always starts with maximal mismatch while infants do not). We could not incorporate the specification of variable hand orientation in ILGM because we did not have access to a quantitative account of the infant hand orientations at reach onset.

The ILGM in the affordance-on condition (modeling the 9-month-old infant) learned to perform grasps more robustly than the affordance-off condition, as the model could associate the affordance (the orientation) of the cylinder with suitable grasp parameters. Figure 10 shows the successful grasping results for the three orientations of the target cylinder. *A1* shows the grasp with supination, whereas *A2* shows a grasp with pronation, both applied to the horizontal cylinder.

Figure 11 shows the plots for difference in hand and cylinder orientations. *A1* and *A2* (Fig. 11) show the differences in hand orientation for the horizontal cylinder trials. The (approximately) constant curve wherein the forearm remains pronated (see Figs. 10, *A2*, and 11, *A2*) corresponds to the power grasp observed in affordance-off condition simulations. The non-constant curve (see Figs. 10, *A1*, and 11, *A1*) describes a reach during which the forearm rotates from a fully pronated position (palm down) to 90° of supination (thumb up) at mid reach, then to the maximum supinated position. To our knowledge, there is no report of whole hand grasping with maximum supination during reaching in infancy, and apparently Lockman et al. (1984) and McCarty et al. (2001) did not observe this kind of grasp either.

In our simulations, supinated and pronated grasps of the horizontal cylinder learned by ILGM (Fig. 11, *A1* and *A2*) were equal in their reward value. One could incorporate a penalty term for energy use in the ILGM reward computation that evaluates how easily a grasp action can be implemented. This would tilt the balance of grasp

Fig. 8 (SE3a) the hand orientation and cylinder orientation difference (error) curves for individual trials demonstrating the deficits that arise when no affordance information is available for the orientation of the target cylinder. The abscissa of the plots represents the normalized time for grasping movements. The columns from left to right correspond to horizontal (0°), diagonal (45°) and vertical (90°) orientation of the cylinder. Upper row (*A1*, *B1*, *C1*): constant class of error curves; lower row (*A2*, *B2*, *C2*): non-constant class of error curves

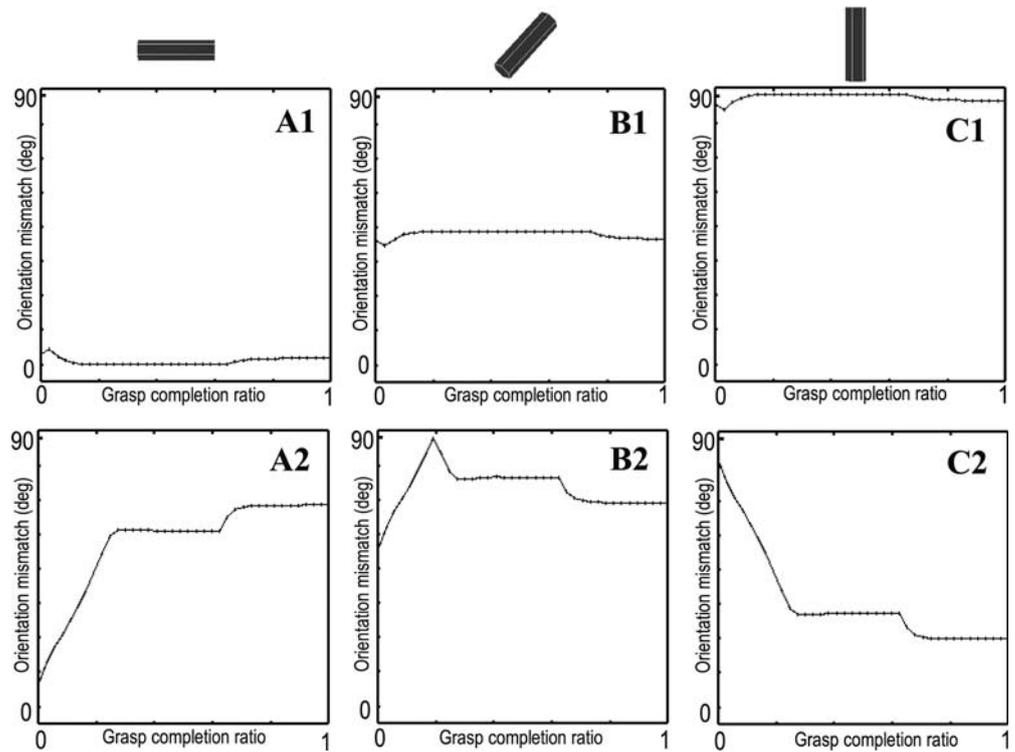


Fig. 9a, b Comparing SE3a, SE3b with infant data. The hand orientation score versus grasp execution is shown. The score indicates extent of mismatch of the hand to the vertical cylinder (0 value indicates perfect match). **a** Dashed line: 5-months-old infant; Solid line: 9-months-old infant (infant data from Lockman et al. (1984)). **b** Dashed line: ILGM in affordance-off (SE3a) case; Solid line: ILGM in affordance-on (SE3b) case

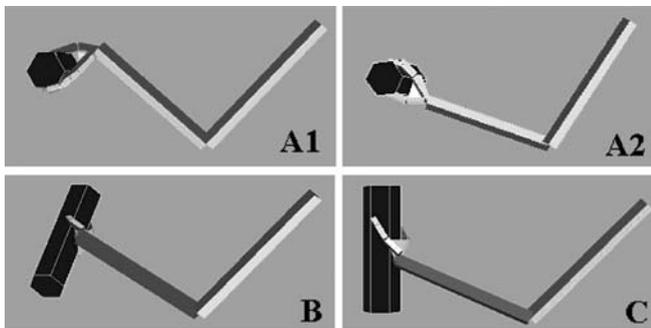
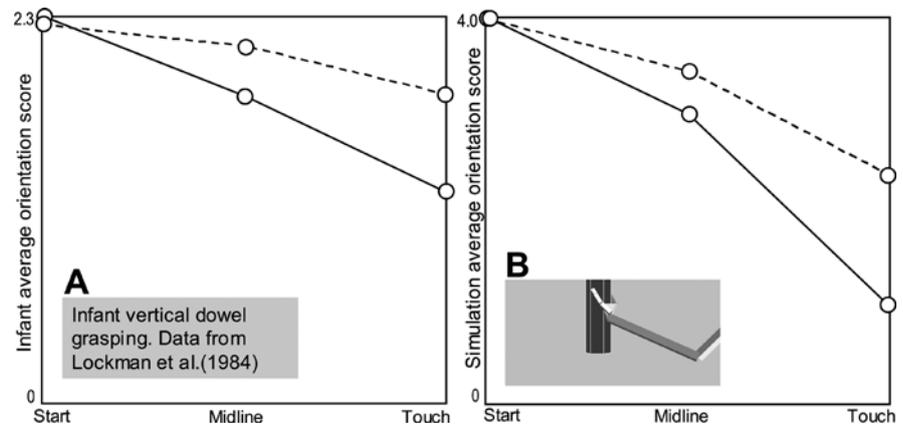


Fig. 10 (SE3b) the grasps performed after ILGM learned the association of hand rotations with the object orientation input (affordance-on condition)

choice towards the pronated grasp (Fig. 11, *A2*) as it requires no forearm supination.

One notable outcome of the simulations is that power grasping with a supinated hand did not appear during the affordance-off condition. Whereas during the affordance-on condition, supinated and pronated reaches were equally successful. From our results we propose the following predictions of infant learning to grasp for testing the ILGM. First, because we have no compelling reason at this time to think infants are physically unable to supinate during arm movements, we predict that neonates and youngest infants employ supinated reaches more frequently than infants 4–6 months of age. Second, we predict that the supinated reaches of neonatal and young infants do not generate successful contacts and rewards. Third, we predict that lack of successful contact extinguishes supinated reaches at 4–6 months of age as other strategies yield greater success rates. Fourth, we predict that all successful reaches at 4–6 months employ a pronated forearm orientation and approach from either

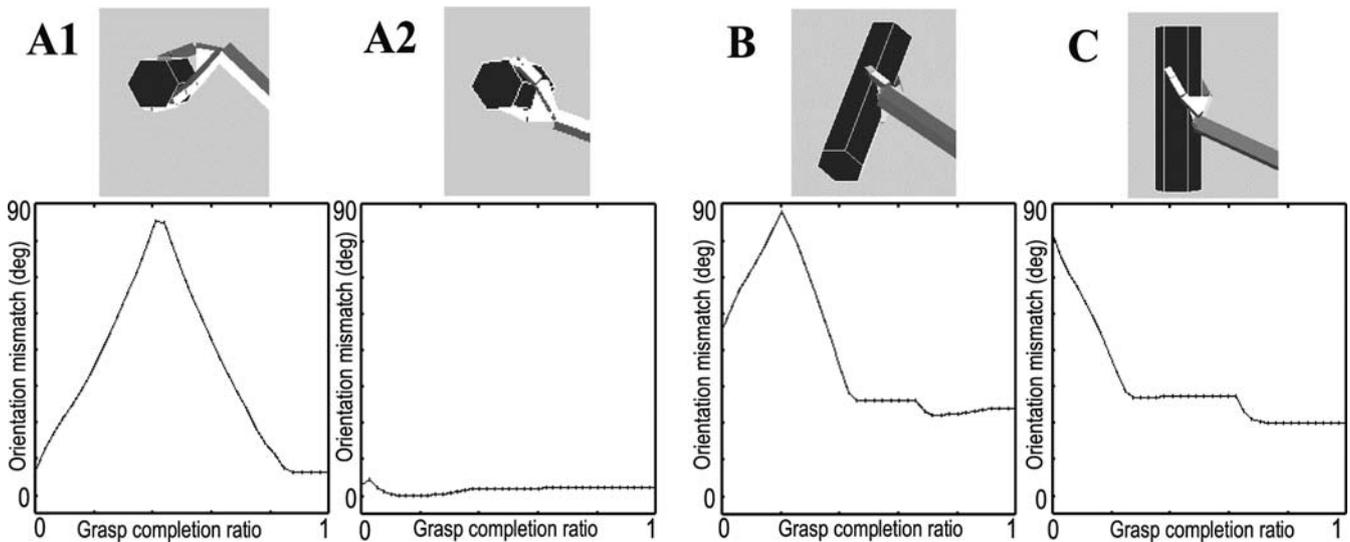


Fig. 11 (SE3b) the grasp errors that ILGM produced in the affordance-on condition. *Left two plots* show typical error curves for the horizontal cylinder with supination (*A1*) and without supination

(*A2*) during a reach to power grasp. *B* shows a typical error curve for the diagonal cylinder, and *C* shows a typical error curve for the vertical cylinder

above or to the side of an object. Fifth, because the simulation was equally successful in completing supinated and pronated reaches in the affordance-on condition, we predict that infants start employing supinated reaches more frequently as they acquire the ability to recognize object affordances. And sixth, we predict that between 7–9 months of age, infants will increasingly supinate the forearm and approach objects from underneath.

SE4: generalization to different target locations

After learning to grasp with varying object locations, the model acquired a grasp planning ability that generalizes to different locations in the workspace. Figure 12a illustrates the learning achieved as superimposed images of completed grasps. Note that the locations of the objects in Fig. 12a were not used in training. When we tested the generalization of the learning (over the centers of the squares formed by four equidistant nodes in Fig. 13a plus the extrapolated locations over the right and bottom grid boundaries) we found that 85/100 grasps conformed to the successful grasp criteria used in training. The lower three grasp actions in Fig. 12 are anatomically hard to achieve but they were included to demonstrate the full range of actions that ILGM learned. The illustration in Fig. 12a only depicts a possible grasp for each location out of many alternatives. We re-ran the trained model on exactly the same object locations as in Fig. 12a. A set of different grasping configurations was generated by ILGM, showing that ILGM can both *generalize* and represent *multiple grasp plans* for each target location (compare Fig. 12a and b).

For each object location, Hand Position layer generates a probability distribution of azimuth, elevation and radius parameters, (α, β, r) , specifying an offset from which the hand will approach the target object. By summing over the

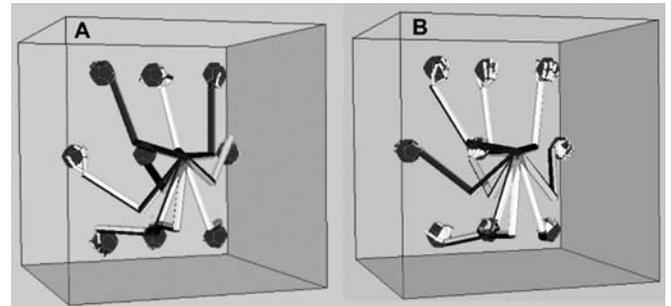


Fig. 12a, b (SE4) the trained ILGM executed grasps twice (*a, b*) to objects located at nine different locations in the workspace. The grasp locations were not used in the training. Note that although the target locations were the same for *a* and *b* cases, the grasps executed differed showing that ILGM learned a *set* of grasps for each target location

radius component, we visualize each object location's 'grasp affordability' as an image map. In Fig. 13b each small image patch is a representation of the (azimuth, elevation) probability distribution, with black having the highest and white having the lowest probability. The correspondence between the object location and the image map matrix is indicated by the common labels 1, 2, 3, 4 in Fig. 13a, b. The label *O* indicates the position of the shoulder of the simulated arm. The results show that the ipsilateral workspace affords a wide range of grasping: the *black strips* in the plots indicate a range of elevation values, i.e., a lower-to-upper range of approaches to the target affords wrist rotations leading to successful grasping. Thus, the results predict that infants are more likely to learn the 'easy locations' (i.e., the locations with darker image maps) that offer large sets of grasping possibilities, early in grasp development.

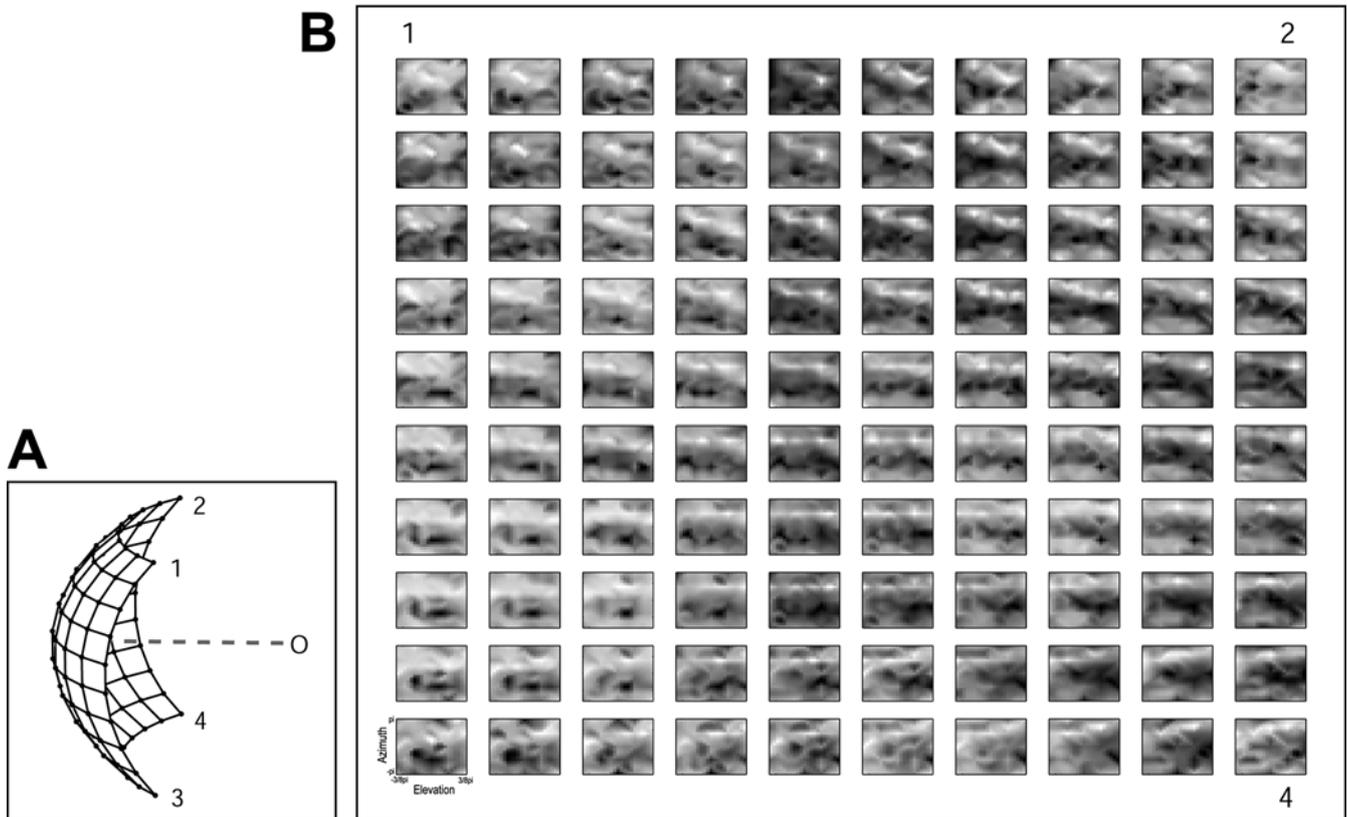


Fig. 13a, b (SE4) **a** The object locations used in training is illustrated. **b** The dependency of the Hand Position layer output on the object location is shown. The grid in the left panel also indicates the object locations used for the generation of the image map matrix

in panel **b**. The topology correspondence is indicated with the common labels 1, 2, 3 and 4. The label *O* indicates the shoulder position

Discussion

This paper presented the Infant Learning to Grasp Model (ILGM) and simulations yielding predictions about the process of motor learning in infancy. We first showed that with just a limited set of behaviors (reaching and grasp reflex) goal-directed trial and error learning yields grasping behavior. We showed that palm-orienting could be acquired through learning, and that object affordance information was not a prerequisite for successful grasp learning. The ‘grasp menu’ acquired via learning included various precision grips in addition to power grasps. Learning to grasp via imitation requires visual analysis of the hand in action and its relation to the goal affordance. Thus, infants’ early grasping could be mediated by self-learning mechanisms rather than by learning by imitation.

During development, infants encounter constraints and find ways to act within the limitations of the environment and reach context. We showed how task constraints could play a role in shaping an infant’s grasping behavior. We simulated a situation where a small cube was placed on, or just under, a table. The model was asked to interact with this simple environment. The grasping configurations learned by the model reflected the task constraints. The model did not acquire whole hand prehension grasps but acquired grasps that reach the cube from above or below,

as permitted by the constraints, using the precision grip to avoid a collision with the table.

Finally, we analyzed what object affordance may add to ILGM by simulating the experimental set up of Lockman et al. (1984). With this simulation we showed not only the improvement in grasp execution, the matching of an infant’s hand orientation with that of a dowel, but also an improvement pattern comparable to that observed by Lockman et al. (1984) and McCarty et al. (2001). These findings indicate that ILGM captured the behavior of 5- and 9-month-old infants via control of affordance availability.

The work allowed us to provide computational support for the following four hypotheses:

- (H1) Early postnatal period infants can acquire the skill to orient their hand towards a target, rather than innately possessing it.
- (H2) Infants are able to acquire grasping skills before they develop an elaborate adult-like object visual analysis capability suggesting that imitation might not play a crucial role for grasp development in early infancy.

- (H3) Task constraints due to environmental context (or the action opportunities afforded by the environment) are a factor shaping infant grasp development.
- (H4) Grasping performance in the ‘affordance-off stage’ mediates the development of a visual affordance extraction circuit.

In various simulations, we encoded the affordances as a population code rather than modeling how affordances are acquired. Here we suggest, as a basis for future simulations, that our ILGM findings, by showing how a repertoire of grasps can be acquired, lay the basis for a model of the development of infants’ further visuomotor skills. In particular, *Affordance learning* can be mediated by the phase we have modeled by ILGM. Successful grasp attempts of an infant categorize the interacted objects into (possibly overlapping) classes depending on the grasps afforded. Then, a compact representation of the object features relevant for manipulation (i.e., affordances) can be obtained by detecting the commonalities within these classes (e.g., shape, orientation, etc.)

Adult grasping studies suggest that reach and grasp components of a grasping action are independently programmed and simultaneously implemented (Jeannerod and Decety 1990). However, in infants the reaching component first dominates action; infants first learn to reach, or how to control projection of the hand, within a given workspace. We suggest that over developmental time a grasp planning inversion takes place, that is, later in infancy the desired grasp configurations begin to alter the control of reach. The inversion is driven by the ‘joy of grasping’ as infants learn from the manipulations afforded by successful reaches so that, eventually, a representation of the desired manipulation of an object (e.g., power grasping) would be an *input* to the reach and grasp planning circuit.

We conclude the paper by relating ILGM to modular reinforcement learning models and presenting future research directions. The existing models of modular or hierarchical reinforcement learning aim at decomposing the task into simpler subtasks so that learning can be accelerated (Dayan and Hinton 1993; Dietterich 2000; Doya et al. 2002). The major problem is how to discover proper subtasks automatically (see Barto and Mahadevan 2003 for a review). In ILGM, we had prior information about the problem and thus implemented a modular system manually, tuned for the task of learning to grasp. Our implementation decomposed the action (i.e., *enclosure rate, hand position, wrist rotation*) space and enforced dependency constraints between action components via the connection matrices.

When experimenting with ILGM, although the stability analysis of a grasp included forces, the hand-object interaction and the gravity were not simulated as a true physical system (i.e., the hand/arm model was kinematics based). Thus, a natural continuation of the study would be to test the ILGM model with a realistic simulation of the dynamics of the hand and arm. This will enable the

simulation of learning with anatomical and physiological constraints (i.e., energy use, anatomical discomfort, etc.), which will better reflect the course of grasp development. In a natural situation after contacting an object, an infant switches to a tactile search phase to compensate for the errors in the open-loop grasp plan. However, tactile search was not possible in the simulation, as it requires a complex model of the fingers to provide, for example, the compliance property of human fingers. In addition, the mechanoreceptors of the fingers help stabilize object by relaying slip and indentation based feedback forming a fast grasp stabilizing motor servo (Johansson and Westling 1987a, b; Westling and Johansson 1987). Thus, we propose that the process of infant grasping can be viewed as the cooperation of three processes: (1) spinal reflex loops for stabilizing a grasped object; (2) tactile search of hand configuration suitable for the target object; and (3) generating open-loop grasp plans that will yield successful grasping when contacted with the target object.

This paper focused on (3) and used the palmar reflex in the simulations. The modeling of mechanoreceptor functioning and spinal reflex loops (1) and tactile search (2) would add to our understanding of development of grasping by showing how tactile input reduces the complexity of grasp learning in cortical regions. From a technical point of view, implementation of components (1) and (2), would allow learning with fewer open-loop grasp trials and leave computational power for detailed physics simulation of the fingers and the object. Another major modeling study would be to move from infant to adult grasping by analyzing the development of visual feedback circuits for manual manipulation and grasping. We propose that grasping capability first (i.e., in infancy) relies on (1), (2) and (3), and then it is augmented by visual feedback control, which is necessary for dexterous hand manipulation. Based on our predictions from the current study, we suggest that the training stimuli for the visual feedback circuit are provided by the open-loop grasping performance. The construction of a biologically realistic model of visual feedback control that bootstraps its function by learning from the visual, proprioceptive and haptic data produced during infant grasp execution mediated and learned by processes (1–3) will present a causally complete model that explains the transition from infant grasping to adult grasping.

Acknowledgements This work was supported in part by a Human Frontier Science Program grant to MAA, which made possible useful discussions with Giacomo Rizzolatti and Vittorio Gallese, to whom we express our gratitude. The work of MAA was supported in part by grant DP0210118 from the Australian Research Council (R. A. Owens, Chief Investigator). The work of EO was supported in part by ATR Computational Neuroscience Laboratories, Kyoto, Japan.

Appendix 1

The simulation loop

The global logic of a simulation session is given below. This applies to all simulations except Simulation 1, where the function of the Wrist Rotation layer output is replaced by the automatic palm orientation described in Appendix 2. The Affordance layer encodes the orientation of the target object in SE3b, and the location of the object in SE4. In the remaining simulations it encodes the existence of a target object.

- Step 1: Encode the affordances presented to the circuit in Affordance layer (A). Set the vector o as the center of the target object.
- Step 2: Compute the outputs of Virtual Finger (V), Hand Position (H) and Hand Rotation (R) layers.
- Step 3: Generate movement parameters v, h, r .
- Step 4: Generate reach using the parameters v, h and r and o while monitoring for contact.
- Step 5: On contact (or failure to contact) compute a contact list.
- Step 6: Compute stability and generate reward signal rs based on the contact list.
- Step 7: Adapt weights using rs .
- Step 8: Go to Step 2, unless interrupted by the user.

Hand Position layer output encoding

The Hand Position layer generates a vector composed of azimuth, elevation and radius (α, β, r). The vector is related to the rectangular coordinates (in a left-handed coordinate system) as follows:

$$x = r \cos \beta \sin \alpha, \quad y = r \sin \beta, \quad z = -r \cos \beta \cos \alpha \quad (17)$$

Figure 14 illustrates the conversion graphically.

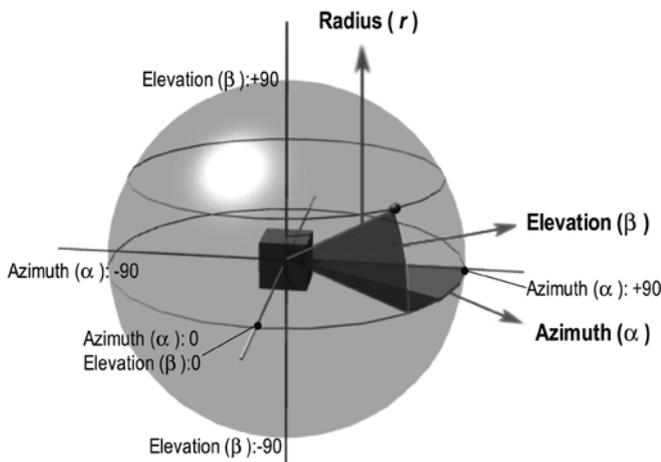


Fig. 14 The conventions used in the spherical representation of the Hand Position layer output. Note that azimuth values of 180 and -180° coincide

Appendix 2

The arm/hand model

Since the arm/hand model we used in our simulations is a kinematics one, the absolute values of the modeled lengths are irrelevant and thus not specified here; however, the relative sizes of the segments are as shown in Fig. 15.

The arm is modeled to have a 3DOF joint at the shoulder to mimic the human shoulder ball joint and a 1DOF joint in the elbow for lower arm extension/flexion movements (see Fig. 15). The wrist is modeled to have 3DOFs to account for extension/flexion, pronation/supination, and ulnar and radial deviation movements of the hand. Each finger except the thumb is modeled to have 2DOFs to simulate metacarpophalangeal and distalinterphalangeal joints of human hand. The thumb is modeled to have 3DOFs, one for the metacarpophalangeal joint and the remaining two for extension/flexion and ulnar and radial extension movements of the thumb (i.e., for the carpometacarpal joint of the thumb).

Inverse kinematics

The simulator coordinate system for the forward and inverse kinematics is left-handed. The zero posture and the effect of positive rotations of the arm joints are shown in Fig. 16.

When we mention the reach component, we imply the computation of trajectories of the arm joint angles ($\theta_1, \theta_2, \theta_3, \theta_4$) to achieve a desired position for the end effector (i.e. the inverse kinematics problem). The end effector could be any point on the hand (e.g., the wrist, index finger tip, middle finger joints, etc.) as long as it is fixed with respect to the arm and the length of the lower limb of the arm is extended (for the sake of kinematics computa-

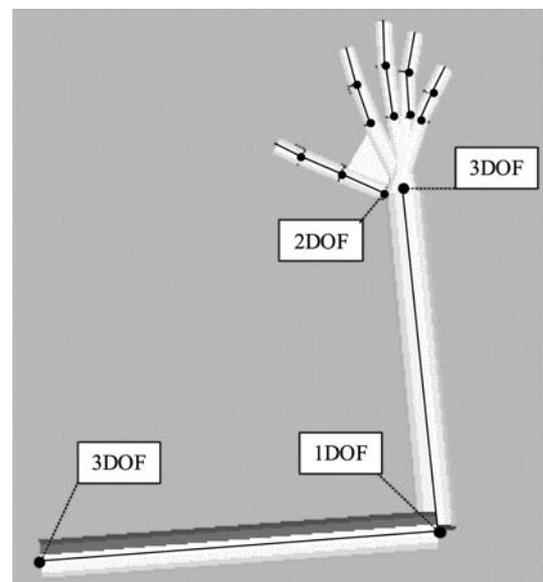


Fig. 15 The arm/hand model has 19 DOFs as shown

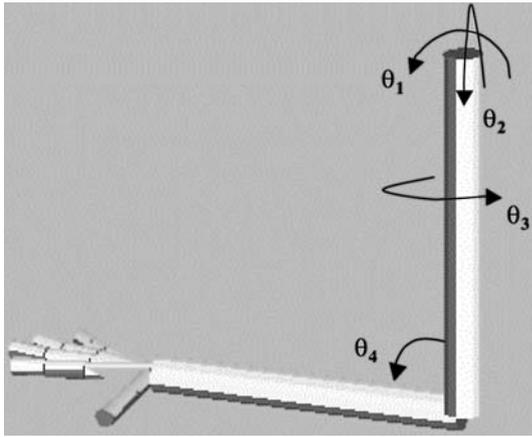
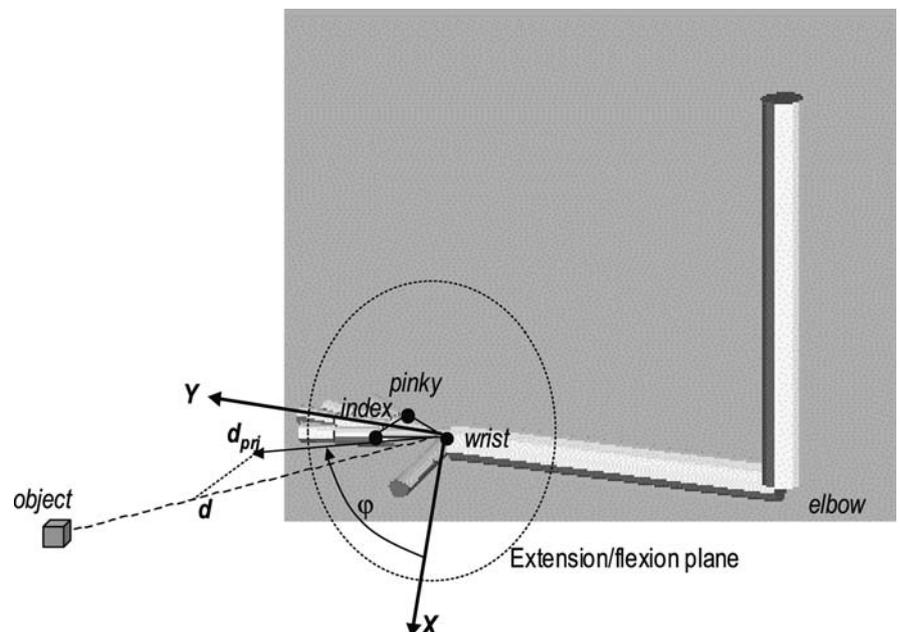


Fig. 16 Zero posture of the arm/hand together with arm DOFs is shown. The arrows indicate the positive rotations

tions) to account for the end effector position. The end effector used in the simulations was the tip of either the index or the middle finger (see Appendix 3). The forward mapping, F , of the kinematics chain is a vector-valued function of the joint angles relating the joint angles to the end-effector position.

$$J = \begin{bmatrix} 0 & l_2(-c_2s_4 + s_2s_3c_4) - l_1c_2 \\ l_2(-s_1c_2s_4 - (s_1s_2s_3c_4 + c_1c_3c_4)) + l_1s_1c_2 & l_2(c_1s_2s_4 + c_2s_3c_1c_4) + l_1c_1c_2 \\ l_2(-c_1c_2s_4 + c_1c_4s_2s_3 - c_3c_4s_1) - l_1c_1c_2 & l_2(-s_1s_2s_4 + c_2c_4s_1s_3) + l_1s_1s_2 \\ -l_2(c_2c_3c_4) & l_2(-s_2c_4 + c_2s_3s_4) \\ l_2(c_1s_2c_3c_4 + s_1s_3c_4) & -l_2(-c_1c_2c_4 - c_1s_2s_3s_4 + s_1c_3s_4) \\ l_2(-c_3c_4s_1s_2 - c_1c_4s_3) & l_2(-s_1c_2c_4 - s_1s_2s_3s_4 - c_1c_3s_4) - l_1c_1c_2 \end{bmatrix} \quad (19)$$

Fig. 17 The automatic hand orientation is modeled as minimizing the angle between the vectors X (palm normal) and d



The Jacobian transpose method for inverse kinematics can be derived as a gradient descent algorithm by minimizing the square of the distance between the current (p) and the desired end-effector position ($p_{desired}$). The key to the algorithm is a special matrix called the geometric Jacobian matrix (J), which relates end-effector Cartesian velocity to the angular velocities of the arm joints (Sciavicco and Siciliano 2000):

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = J \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \\ \dot{\theta}_4 \end{bmatrix} \quad (18)$$

or in vector notation $\dot{p} = J\dot{\theta}$.

Representing the upper arm length and the (extended) lower arm length with l_1 and l_2 , respectively; and abbreviating $\sin(\theta_1)$ and $\cos(\theta_1)$ with s_1 and c_1 ; $\sin(\theta_2)$ and $\cos(\theta_2)$ with s_2 and c_2 ; $\sin(\theta_3)$ and $\cos(\theta_3)$ with s_3 and c_3 ; $\sin(\theta_4)$ and $\cos(\theta_4)$ with s_4 and c_4 , the Jacobian matrix of our arm model can be written as:

Then the algorithm is simply composed of iterating the following update rule until $p=p_{desired}$, where η represents

the joint angle update rate.

$$\theta_{t+1} = \theta_t + \eta J_{\theta_t}^T (p_{desired} - p_t) \quad (20)$$

Note that the time dependency of the variables and the dependency of the Jacobian matrix on the joint angles are indicated with subscripts.

Automatic hand orientation

The automatic hand orientation employed in SE1a is modeled as minimizing the angle between the palm normal (X) and the vector (d) connecting the center of the object to the index finger knuckle using the hand's extension/flexion degree of freedom (see Fig. 17). The angle is minimized when the palm normal coincides with the projection of d on to the extension/flexion plane of the hand.

When the hand makes a rotation of ϕ radians as illustrated in Fig. 17, the palm normal coincides with the projection of d on to the extension/flexion plane of the hand. Noting that *object*, *index*, *pinky*, *wrist* and *elbow* in Fig. 17 represent three-dimensional position vectors, the angle ϕ can be obtained as follows:

$$X = \frac{(pinky - index) \otimes (wrist - index)}{|(pinky - index) \otimes (wrist - index)|} \quad (21)$$

$$Y = \frac{wrist - elbow}{|wrist - elbow|}$$

$$d = object - index \quad (22)$$

$$d_{prj} = \langle X \cdot d \rangle X + \langle Y \cdot d \rangle Y$$

$$\phi = -\cos^{-1} \left(\frac{\langle d_{prj} \cdot X \rangle}{|d_{prj}|} \right) \quad (23)$$

When automatic hand orientation is engaged, the hand is rotated by ϕ radians at each cycle of the simulation while reach is taking place. Note that when d_{prj} is zero, the angle ϕ is not defined. In that case, the angle is returned as zero (i.e., hand is not rotated). This situation happens when the extension/flexion movement of the hand has no effect on the angle between the palm normal and d ; in other words, when d is vertical to both X and Y .

Appendix 3

ILGM simulation parameters

The main behavior of the simulation system is determined by a resource file where many simulation parameters can be set. In this file, the three-dimensional positions and vectors are defined using a spherical coordinate system.

The PAR, MER and RAD tags are used to indicate elevation, azimuth and radius components, respectively.

Object axis orientation range parameters

Object axis orientation range parameters define the minimum and maximum allowed tilt of the object around the z-axis (in the frontal plane). Ten units are allocated for encoding the tilt amount. These parameters are only used in SE3.

$$\begin{aligned} \text{minTILT} & 0 \\ \text{maxTILT} & 90 \end{aligned} \quad (24)$$

Base learning rate parameter

Base learning rate parameter is used as the common multiplier for all the learning rates in the grasp learning circuit as $\eta_{AR} = \eta_{VR} = \eta_{HR} = \text{eta}$; $\eta_{AV} = \eta_{AH} = \eta/\text{MAXROTATE}$.

$$\text{eta} (\eta) 0.5 \quad (25)$$

LGM layer size parameters

LGM layer size parameters define the number of units to allocate for the layers generating the motor parameters. The Virtual finger (V) layer is composed of ten units which specify the synergistic control of the fingers. In what follows BANK, PITCH and HEADING tags indicate the supination-pronation, wrist extension-flexion and radial/ulnar deviation movements, respectively. The size of the Wrist Rotation (R) layer is determined with the following parameters (in this example $9 \times 9 \times 1$ units will be allocated):

$$\begin{aligned} \text{hand_rotBANK_code_len} & 9 \\ \text{hand_rotPITCH_code_len} & 9 \\ \text{hand_rotHEADING_code_len} & 1 \end{aligned} \quad (26)$$

In what follows, the tags locMER, locPAR and locRAD indicate the Hand Position (H) layer components. The size of the Hand Position layer is determined with the following parameters (in this example $7 \times 7 \times 1$ units will be allocated)

$$\begin{aligned} \text{hand_locMER_code_len} & 7 \\ \text{hand_locPAR_code_len} & 7 \\ \text{hand_locRAD_code_len} & 1 \end{aligned} \quad (27)$$

Learning session parameters

Learning session parameters define the behavior of the simulator during learning. For a learning session, the simulator makes MAXBABBLE number of reach/grasp attempts. For each approach-direction (H layer output), the simulator makes MAXROTATE grasping attempts. After MAXREACH reaches are done, the next input condition is selected (e.g., the object orientation is changed). MAXBABBLE limits the maximum number of attempts the simulator will make. A particular simulation may be stopped at any instant. The saved connection weights then can be used for testing the performance later. Reach2Target parameter indicates which part of the hand should be used as the end effector by MG module for reach execution. The possible values are [INDEX, MIDDLE, THUMB] \times [0,1,2] where 2 indicates the tip and 0 indicates the knuckle. An example set of parameter specification is as follows:

MAXREACH	5	
MAXROTATE	7	
MAXBABBLE	10000	(28)
weightSave	4500	
Reach2Target	INDEX1	

Grasp stability parameters

Grasp stability parameters define the acceptable grasps in terms of physical stability. costThreshold specifies the allowable inaccuracy in grasping. Ideally, the cost of grasping should be small indicating that the grasp is successful. Empirically, a threshold ($E_{threshold}$) value between 0.5 and 0.8 gives a good result for the implemented cost function. If the distance of the touched object to the palm is less than palmThreshold and the movement of the object due to finger contact is towards the palm, then the palm is used as a virtual finger to counteract the force exerted by the fingers. The negReinforcement parameter specifies the level of punishment returned when a grasp attempt fails (rs_{neg}). Empirically values greater than -0.1 and less than 0 result in good learning. Generally, a large negative reinforcement overwhelms the positively reinforced plans before they have chance to get represented in the layers.

costThreshold	0.8	
palmThreshold	150	(29)
negReinforcement	-0.05	

Exploration and exploitation parameter

α (Randomness) specifies how often to use the learned distribution to generate grasp plans. A value of 1 means always use random parameter selection, while a value of 0

means always generate parameters from the current distribution of the layer. In all the simulations, the Virtual finger layer used the probability distribution representation to generate enclosure parameter (v). Next, we present the parameters used for other layers in the simulation experiments. The default parameters values used as examples in the descriptions above are not repeated here (Tables

Parameters for SE1a

hand_rotBANK_code_len	N/A
hand_rotPITCH_code_len	N/A
hand_rotHEADING_code_len	N/A
hand_locMER_code_len	10
hand_locPAR_code_len	10
hand_locRAD_code_len	10
MAXREACH	N/A
MAXROTATE	N/A
MAXBABBLE	1000
Reach2Target	INDEX0
costThreshold ($E_{threshold}$)	0.75
palmThreshold	125
negReinforcement (rs_{neg})	-0.1
Randomness (α)	0.85

a,

Parameters for SE2

hand_rotBANK_code_len	9
hand_rotPITCH_code_len	9
hand_rotHEADING_code_len	1
hand_locMER_code_len	7
hand_locPAR_code_len	7
hand_locRAD_code_len	1
MAXREACH	1
MAXROTATE	7
MAXBABBLE	45000
Reach2Target	INDEX2
costThreshold ($E_{threshold}$)	0.80
palmThreshold	150
negReinforcement (rs_{neg})	-0.05
Randomness (α)	1

b,

Parameters for SE1b

hand_rotBANK_code_len	11
hand_rotPITCH_code_len	11
hand_rotHEADING_code_len	6
hand_locMER_code_len	6
hand_locPAR_code_len	6
hand_locRAD_code_len	1
MAXREACH	N/A
MAXROTATE	55
MAXBABBLE	10000
Reach2Target	INDEX0
costThreshold ($E_{threshold}$)	0.75
palmThreshold	125
negReinforcement (rs_{neg})	-0.1
Randomness (α)	0.95

c,

Parameters for SE3

hand_rotBANK_code_len	12
hand_rotPITCH_code_len	7

Parameters for SE3

hand_rotHEADING_code_len	1
hand_locMER_code_len	5
hand_locPAR_code_len	5
hand_locRAD_code_len	5
MAXREACH	1
MAXROTATE	25
MAXBABBLE	20000
Reach2Target	MIDDLE0
costThreshold ($E_{threshold}$)	0.85
palmThreshold	150
negReinforcement (rs_{neg})	-0.1
Randomness (α)	0.95

d and

Parameters for SE4

hand_rotBANK_code_len	10
hand_rotPITCH_code_len	7
hand_rotHEADING_code_len	5
hand_locMER_code_len	10
hand_locPAR_code_len	10
hand_locRAD_code_len	1
MAXREACH	10
MAXROTATE	30
MAXBABBLE	10000
Reach2Target	MIDDLE0
costThreshold ($E_{threshold}$)	0.75
palmThreshold	128
negReinforcement (rs_{neg})	-0.1
Randomness (α)	1.0

e).

References

- Arbib MA, Hoff B (1994) Trends in neural modeling for reach to grasp. In: Bennett KMB, Castiello U (eds) Insights into the reach to grasp movement. North-Holland, Amsterdam, New York, pp xxiii, 393
- Barto A, Mahadevan S (2003) Recent advances in hierarchical reinforcement learning discrete-event systems. (in press)
- Baud-Bovy G, Soechting JF (2001) Two virtual fingers in the control of the tripod grasp. *J Neurophysiol* 86:604–615
- Bayley N (1936) The California infant scale of motor development, birth to three years. University of California Press, Berkeley, CA
- Bernstein NA (1967) The coordination and regulation of movements. Pergamon Press, Oxford
- Berthier NE, Clifton RK, Gullapalli V, McCall DD, Robin DJ (1996) Visual information and object size in the control of reaching. *J Motor Behav* 28:187–197
- Bradley NS (2000) Motor control: developmental aspects of motor control in skill acquisition. In: Campbell SK, Van der Linden DW, Palisano RJ (eds) Physical therapy for children, a comprehensive reference for pediatric practice, 2nd edn. Saunders, Philadelphia, pp 45–87
- Butterworth G, Verweij E, Hopkins B (1997) The development of prehension in infants: Halverson revisited. *Brit J Dev Psychol* 15:223–236
- Clifton RK, Muir DW, Ashmead DH, Clarkson MG (1993) Is visually guided reaching in early infancy a myth. *Child Dev* 64:1099–1110
- Dayan P, Hinton G (1993) Feudal reinforcement learning. In: Lippman DS, Moody JE, Touretzky DS (eds) Advances in neural information processing systems, San Mateo, CA, pp 5:271–278
- Diamond A, Lee EY (2000) Inability of five-month-old infants to retrieve a contiguous object: a failure of conceptual understanding or of control of action? *Child Dev* 71:1477–1494
- Dietterich TG (2000) Hierarchical reinforcement learning with the MAXQ value function decomposition. *Journal of Artificial Intelligence Research* 13:227–303
- Doya K, Samejima K, Katagiri K, Kawato M (2002) Multiple model-based reinforcement learning. *Neural Comput* 14:1347–1369
- Fagg AH, Arbib MA (1998) Modeling parietal-premotor interactions in primate control of grasping. *Neural Netw* 11:1277–1303
- Fearing RS (1986) Simplified grasping and manipulation with dexterous robot hands. *IEEE Journal of Robotics and Automation* 2:188–195
- Gibson EJ (1969) Principles of perceptual learning and development. Prentice-Hall, Englewood Cliffs, NJ
- Gibson EJ (1988) Exploratory behavior in the development of perceiving, acting and acquiring of knowledge. *Ann Rev Psychol* 39:1–41
- Halverson HM (1931) An experimental study of prehension in infants by means of systematic cinema records. *Genet Psychol Monogr* 10:107–285
- Iberall T, Arbib MA (1990) Schemas for the control of hand movements: an essay on cortical localization. In: M.A. G (ed) Vision and action: the control of grasping. Ablex, Norwood, NJ
- Jeannerod M, Decety J (1990) The accuracy of visuomotor transformation: an investigation into the mechanisms of visual recognition of objects. In: Goodale MA (ed) Vision and action: the control of grasping. Ablex Pub. Corp., Norwood, NJ, pp viii, 367
- Jeannerod M, Arbib MA, Rizzolatti G, Sakata H (1995) Grasping objects—the cortical mechanisms of visuomotor transformation. *Trends Neurosci* 18:314–320
- Jeannerod M, Paulignan Y, Weiss P (1998) Grasping an object: one movement, several components. *Novartis Foundation Symposium* 218:5–16; discussion 16–20
- Johansson RS, Westling G (1987a) Signals in tactile afferents from the fingers eliciting adaptive motor responses during precision grip. *Exp Brain Res* 66:141–154
- Johansson RS, Westling G (1987b) Significance of cutaneous input for precise hand movements. *Electroencephalogr Clin Neurophysiol* 39[Suppl]:53–57
- Lantz C, Melen K, Forsberg H (1996) Early infant grasping involves radial fingers. *Dev Med Child Neurol* 38:668–674
- Lasky RE (1977) The effect of visual feedback of the hand on the reaching and retrieval behavior of young infants. *Child Dev* 48:112–117
- Lockman J, Ashmead DH, Bushnell EW (1984) The development of anticipatory hand orientation during infancy. *J Exp Child Psychol* 37:176–186
- MacKenzie CL, Iberall T (1994) The grasping hand. North-Holland, Amsterdam, New York
- McCarty MK, Clifton RK, Ashmead DH, Lee P, Goulet N (2001) How infants use vision for grasping objects. *Child Dev* 72:973–987
- Meltzoff AN (1988) Infant imitation after a 1-week delay: long-term memory for novel acts and multiple stimuli. *Dev Psychol* 24:470–476
- Meltzoff AN, Moore K (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78
- Murata A, Gallese V, Kaseda M, Sakata H (1996) Parietal neurons related to memory-guided hand manipulation. *J Neurophysiol* 75:2180–2186
- Murata A, Gallese V, Luppino G, Kaseda M, Sakata H (2000) Selectivity for the shape, size, and orientation of objects for grasping in neurons of monkey parietal area AIP. *J Neurophysiol* 83:2580–2601
- Newell KM (1986) Motor development in children: aspects of coordination and control. In: Wade MG, Whiting HTA (eds) Motor development in children: aspects of coordination and control. Nijhoff, Boston, pp 341–360

- Newell KM, Scully DM, McDonald PV, Baillargeon R (1989) Task constraints and infant grip configurations. *Dev Psychobiol* 22:817–831
- Newell KM, McDonald PV, Baillargeon R (1993) Body scale and infant grip configurations. *Dev Psychobiol* 26:195–205
- Olmos M, Carranza JA, Ato M (2000) Force-related information and exploratory behavior in infancy. *Infant Behav Dev* 23:407–419
- Oztop E, Arbib MA (2002) 2002, schema design and implementation of the grasp-related mirror neuron system, *Biological Cybernetics* 87:116–140
- Rochat P (1998) Self-perception and action in infancy. *Exp Brain Res* 123:102–109
- Rochat P, Morgan R (1995) Spatial determinants in the perception of self-produced leg movements by 3–5 month-old infants. *Dev Psychol* 31:626–636
- Sakata H, Taira M, Kusunoki M, Murata A, Tanaka Y, Tsutsui K (1998) Neural coding of 3D features of objects for hand action in the parietal cortex of the monkey. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 353:1363–1373
- Sakata H, Taira M, Kusunoki M, Murata A, Tsutsui K, Tanaka Y, Shein WN, Miyashita Y (1999) Neural representation of three-dimensional features of manipulation objects with stereopsis. *Exp Brain Res* 128:160–169
- Sciavicco L, Siciliano B (2000) *Modelling and control of robot manipulators*. Springer, London New York
- Smeets JB, Brenner E (1999) A new view on grasping. *Motor Control* 3:237–271
- Smeets JB, Brenner E (2001) Independent movements of the digits in grasping. *Exp Brain Res* 139:92–100
- Sporns O, Edelman GM (1993) Solving Bernstein's problem: a proposal for the development of coordinated movement by selection. *Child Dev* 64:960–981
- Streri A (1993) *Seeing, reaching, touching: the relations between vision and touch in infancy*. Harvester Wheatsheaf, London, New York
- Sutton RS, Barto AG (1998) *Reinforcement learning: an introduction*. MIT Press, Cambridge, MA
- Taira M, Mine S, Georgopoulos AP, Murata A, Sakata H (1990) Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp Brain Res* 83:29–36
- Thelen E (2000) Motor development as foundation and future of developmental psychology. *Int J Behav Dev* 24:385–397
- Twitchell TE (1970) Reflex mechanisms and the development of prehension. In: Connolly KJ (ed) *Mechanisms of motor skill development*. Academic Press, London, New York
- van Rullen R, Thorpe SJ (2001) Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput* 13:1255–1283
- von Hofsten C (1982) Eye-hand coordination in the newborn. *Dev Psychol* 18:450–461
- von Hofsten C (1984) Developmental changes in the organization of prereaching movements. *Dev Psychol* 20:378–388
- von Hofsten C (1991) Structuring of early reaching movements: a longitudinal study. *J Mot Behav* 23:280–292
- von Hofsten C (1993) The structuring of neonatal arm movements. *Child Dev* 64:1046–1057
- von Hofsten C, Fazel-Zandy S (1984) Development of visually guided hand orientation in reaching. *J Exp Child Psychol* 38:208–219
- von Hofsten C, Ronnqvist L (1988) Preparation for grasping an object: a developmental study. *J Exp Psychol Hum Percept Perform* 14:610–621
- Westling G, Johansson RS (1987) Responses in glabrous skin mechanoreceptors during precision grip in humans. *Exp Brain Res* 66:128–14
- Williams RJ (1992) Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning* 8:229–256
- Zemel RS, Dayan P, Pouget A (1998) Probabilistic interpretation of population codes. *Neural Comput* 10:403–430