Ugur, E., Nagai, Y., Oztop, E., and Asada, M. (Eds) Proceedings of Humanoids 2012 Workshop on Developmental Robotics: Can developmental robotics yield human-like cognitive abilities? November 29, 2012. Osaka, Japan

Goal Babbling: a New Concept for Early Sensorimotor Exploration

Matthias Rolf, Jochen J. Steil

I. Coordination Problems

The human body possesses more than 600 skeletal muscles [1]. Performing purposeful actions to achieve some behavioral goal requires a high degree of coordination of these many degrees of freedom. Yet, human infants are born without the most basic coordination skills like reaching for an object [2], which poses the *learning* of sensorimotor coordination as a fundamental problem in human development. Understanding this ability to learn, and utilizing it for modern robotics systems is one of the major goals of the research fields of *cognitive* [3] and *developmental robotics* [4], [5].

We investigate the learning of reaching skills as an exemplary coordination skill. The problem of reaching is to find motor commands (e.g. joint angles of a robot arm) that move the hand, or the robot's end-effector towards some desired position in space. Thereby motor commands q and outcomes x are connected by a causal relation which is denoted as the forward function $f(q) = x$. Learning needs to invert this relation in order achieve some desired outcome x^*. This problem setup is not only illustrative, but very *prototypical* for other coordination problems: it asks the very general question of *how* to achieve some behavioral goals by means of actions. The skill of reaching itself is also *fundamental* for both robots and humans, since the positioning in space is necessary for any use of the robot's gripper or the human's hand.

Successful reaching skills can be well understood with the notion of *internal models* [6], [7], whereas forward models predict the outcome of an action and inverse models suggest actions in order to achieve a desired outcome. The bootstrapping of internal models without explicit prior-knowledge requires experience that has to be generated by *exploration*. Machine learning approaches thereby traditionally rely on an exhaustive exploration of all possible motor commands, frequently generated by means of an entire random procedure, which is referred to as "motor babbling" [8], [9]. After the data generation phase, learning and coordination can be phrased in a variety of ways [10], [11], [12]. Yet, exhaustive exploration can not be achieved on high-dimensional motor systems such as the human body, modern humanoid robots, or biomimetic robots like elephant trunks. The sheer number of combinations of commands for different actuators is too large to be explored in the lifetime of any learning agent. Understanding human motor development, as well as the successful application of future robotic systems like the *Bionic Handling Assistant* (see Fig. 3), demands for concepts and methods that succeed in sensorimotor learning even without fully exploring the space of possible motor commands.

Matthias Rolf and Jochen J. Steil are with the Research Institute for Cognition and Robotics (CoR-Lab) at Bielefeld University, Germany. Mail: {mrolf,jsteil}@CoR-Lab.Uni-Bielefeld.de

II. Infant Development

The standard models for the learning of coordination skills demand either an exhaustive exploration of all actions [8], [10], [9], [11], [12], or prior knowledge about the action space and forward function [13], [14], [15]. Therefore the acquisition of the coordination skill is divided into separate stages of (random) exploration, learning, and exploitation of the learned mechanisms. However, exhaustive exploration does neither provide an explanation of infant's efficiency in sensorimotor development, nor does it provide a feasible approach for artificial agents to learn in high-dimensional domains.

Nevertheless, the generation of random actions by means of motor babbling has been repeatedly motivated [16], [8], [17] by Piaget's view on infant development [18]. Piaget suggested that development is organized in distinct stages and that, at first, infants do not perform purposeful actions. "The implication of [Piaget's] proposal is that the early behavior of the neonate is essentially random and insensitive to contextual information. Recent research suggests that some re-thinking of this extreme position is necessary" [19]. Contrary to Piaget's suggestions, and the random motor babbling approach, infant developmental studies over the last three decades have found conclusive evidence for coordinated behavior even in newborns. Examples include orienting towards sounds [20], tracking of visual targets [21], and apparent reflexes that have been re-discovered as goal-directed actions [22], [23]. "These behaviors are fragile and inconsistent, which explains why they were overlooked for quite some time" [19].

In the case of reaching, it has been shown that newborns attempt goal-directed movements already few days after birth [24], [25]. Von Hofsten showed that, when salient objects are in the visual field, infants produce more arm movements towards that object, than movements away from it. This indicates a strong role of "learning by doing" instead of random exploration and that infants learn to reach by trying to reach: "Before infants master reaching, they spend hours and hours trying to get the hand to an object in spite of the fact that they will fail, at least to begin with" [26]. From a machine learning point of view, these findings motivate to devise methods that closely intertwine exploration, learning, and exploitation, instead of organizing these aspects in distinct and subsequent stages.

Findings of early goal-directed actions are complemented by studies investigating the structure of infants' reaching attempts over the course of development. When infants perform the first *successful* reaching movements around the age of four months, these movements are controlled in an entire *feedforward* manner [27], [28]. This strongly indicates the use of an inverse model for motor control, which selects one solution and applies it without corrections. The importance of feedforward control does not diminish over the course of development, which is well known from prism-glass experiments [29], but the skill is later on augmented by mechanisms that allow for more adaptive movements and error corrections by means of visual feedback [30]. Moreover, the earliest reaching movements are rather jerky and suboptimal in the sense that the distribution and timing of muscular forces is more complicated than actually necessary [31], [2], [32].

In short, infants appear to follow a very efficient pathway, on which one initial solution is learned, and directly used for goal-directed behavior. Only later on these movements are gradually optimized and become more adaptive. While this pathway is very intuitive, it is orthogonal to the motor-babbling approach which first attempts to gather full knowledge about the sensorimotor space, from which particular solutions can be derived afterwards.

III. A New Concept: Goal Babbling

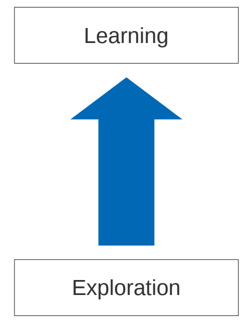
The general idea that connects early goal-directed movements and initial feedforward control is to take redundancy as an *opportunity* to reduce the demand for exploration, instead of a burden that has to be dealt with. If there are multiple ways to achieve some behavioral goal, there is no inherent need to know all of them. Of course, this requires an exploration mechanism that can generate relevant training data without exhaustive exploration. Our main hypothesis is that infants' early goal-directed movements do not only reflect an early exploitation of knowledge, but that they constitute the very mechanism to *generate* that knowledge by exploration, and therefore enable an efficient learning of valid solutions for the coordination problem. Consequently, our main research goal concerns the general mechanism of goal-directed exploration:

> **Research goal 1**: Conceptualize and understand early goal-directed movements as mechanism for the bootstrapping of coordination skills.

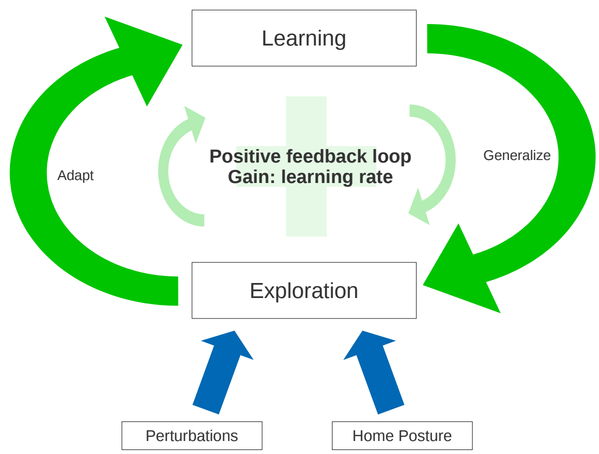
As a basis for this investigation, we have introduced the notion of "goal babbling" [33]:

> **Definition: Goal babbling** is the bootstrapping of a coordination skill by repetitively trying to accomplish multiple goals related to that skill.

A central aspect is, of course, trying to accomplish goals, which corresponds to infants' attempts to perform goal-directed movements. Several other aspects of this definition need to be highlighted in order to distinguish this concept from other approaches: Goal babbling aims at the *bootstrapping* of coordination skills such that a skill can be learned without prior knowledge, or non-goal-directed prior exploration. Goal babbling defines this as a *repeated* process, which implies that the skill acquisition is incremental and ongoing, as opposed to stage-like organizations of exploration and learning [8], [9]. Goal babbling applies to domains with multiple related goals. For reaching problems this is naturally given by a



(a) Uninformed motor babbling



(b) Goal babbling

Fig. 1. In contrast to uninformed exploration processes like motor babbling, exploration and learning mutually inform each other in goal babbling. This organization constitutes a positive feedback-loop during bootstrapping which substantially accelerates learning.

continuous space of possible hand positions. This exploration across multiple goals stands in contrast to typical scenarios in reinforcement learning, in which only a single desired behavior is considered [34]. Given this research goal and the definition of goal babbling, we address several conceptual questions:

- Is goal babbling *possible* at all, and what are the mechanisms necessary to enable it?
- Does it actually permit a bootstrapping that is *scalable to high dimensions*?
- What are observable characteristics of such a bootstrapping process?

IV. Approach and Results

Goal babbling does not refer to a particular algorithm, but to a concept that can be methodically investigated by various means. A recent approach that is compatible with the concept of goal babbling has been introduced in [35]. Baranes' model attempts to learn a partial *forward model*. In this scenario, goal-directed movements are performed by analytically inverting the iteratively learned forward model and performing conventional feedback control. Goal babbling then generates a distribution of actions that lies in the typical regime

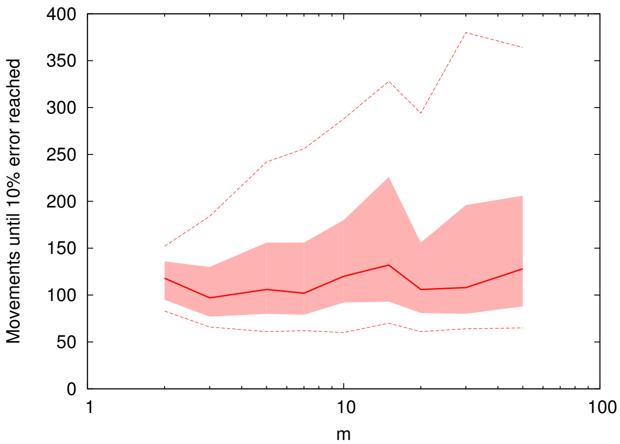


Fig. 2. Goal babbling scales to very high-dimensional problems, as shown by the only marginal increase of exploratory cost for reaching with between $m=2$ and $m=50$ degrees of freedom.

Fig. 3. Goal babbling allows to efficiently learn reaching with the *Bionic Handling Assistant*. The feedforward control with an inverse model allows to cope with intense sensory noise and delays.

of the feedback controller when trying to reach for goals, such that not the entire action space needs to be explored.

In contrast to Baranes' model, we investigate the learning of *inverse models* by means of goal babbling, and therefore focus on learning the coordination skill directly, without relying on analytical inversion mechanisms. This approach resembles infants' developmental pathway, which serves as an example of efficiency, by acquiring at first one valid solution that can be used for feedforward control. Learning inverse models in high-dimensional, redundant domains has, so far, only been possible with error-based mechanisms [13], [14] that use prior knowledge in order to generate a learning signal. The demand for a bootstrapping mechanism clearly disqualifies error-based methods due to their inherent need for prior knowledge. Instead, we focus on learning from autonomously self-generated examples. Learning inverse models from examples was believed to be impossible due to the non-convex solution sets in non-linear redundant domains [14]. Consequently, the second research goal concerns this methodological aspect:

> **Research goal 2**: Enable the learning of inverse models from examples in non-linear and redundant domains.

Finding an exploration scheme that can realize this goal clearly needs to cope with non-convex solution sets. Previous studies have only shown how to deal with non-convexity locally, either by reformulating the problem into a differential one [10], or by using prior knowledge to start learning from a well-initialized state [36]. We show that goal babbling provides an elegant account [33] for this long-standing problem, and demonstrate an algorithm that can utilize *goals as reference structure* in order to resolve possibly inconsistent solutions. However, non-convexity is not the only problem to deal with. While non-convexity makes it difficult to handle multiple solutions for the same outcome, the initial problem is to *find at least one correct solution* to realize the desired outcomes and, hence, to invert the causal relation of the forward function in a reliable manner. This *inversion of causality* is a general problem for exploration schemes, since the direction $x \rightarrow q$ can not be directly queried within coordination problems. Random motor babbling can theoretically solve the problem because it simply explores all actions, such that the necessary ones are also explored. This, however, is practically not feasible in high-dimensional domains. The inversion of causality has a distinct characteristic in goal-directed exploration schemes which tend to get stuck in only partial solutions of the coordination problem [37], [38], [39], in which only a subset of goals can be successfully reached. The general pattern to solve that problem is to introduce exploratory noise into the process [40], [36]. We show novel results for a mathematical theory of example-based learning of inverse models, and provide a proof [41] that goal babbling succeeds in linear domains.

Given a goal babbling method that can learn inverse models from examples, the consequential goal is to practically prove the success and usefulness of goal babbling in high-dimensional domains:

> **Research goal 3**: Devise a practical algorithm for goal babbling that is scalable, fast, and applicable in real-world scenarios.

We demonstrate an online goal babbling algorithm [42] and show that the method does indeed permit enormous scalability (see Fig. 2). This can be achieved because goal-directed exploration allows to leave out redundant choices of actions as soon as there are known ways that solve the problem. The experiments point out that goal babbling constitutes a positive feedback loop during bootstrapping, in which exploration and learning reinforce each other. This positive feedback loop is identified as an important conceptual property of goal babbling that is in line with the *dynamic systems perspective* [43] on development. Experiments demonstrate that it allows to achieve human-level [44] learning speed. We finally demonstrate the practical use of the approach to learn the inverse kinematics of the *Bionic Handling Assistant* (see Fig. 3), which is a relevant and very challenging use case [45].

The results demonstrate the theoretical as well as practical validity of our algorithmic approach. It thereby provides a coherent and constructive explanation of how infants' early goal-directed movements might lead to their rapid initial mastery of feedforward-controlled reaching movements. Other implementations of goal babbling have recently been proposed, and confirm the success of goal babbling, as well as its superiority over motor babbling in terms of bootstrapping efficiency [46], [47], [48]. These results also demonstrate the general validity of the goal babbling concept, which provides a new framework to foster research on infant development as well as robotic systems.

REFERENCES

[1] J. P. Welsh and R. Llins, "Some organizing principles for the control of movement based on olivocerebellar physiology," *Progress in Brain Research*, vol. 114, pp. 449–461, 1997.

[2] J. Konczak, M. Borutta, and J. Dichgans, "The development of goal-directed reaching in infants ii. learning to produce task-adequate patterns of joint torque," *Experimental Brain Research*, vol. 113, no. 3, pp. 465–474, 1997.

[3] S. Kopp and J. J. Steil, "Special corner on cognitive robotics," *Cognitive Processing*, vol. 12, no. 4, pp. 317–318, 2011.

[4] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2003.

[5] M. Asada, K. Hosoda, Y. Kuniyoshi, H. Ishiguro, T. Inui, Y. Yoshikawa, M. Ogino, and C. Yoshida, "Cognitive developmental robotics: A survey," *IEEE Transactions on Autonomous Mental Development*, vol. 1, no. 1, pp. 12–34, 2009.

[6] D. M. Wolpert, R. C. Miall, and M. Kawato, "Internal models in the cerebellum," *Trends in Cognitive Science*, vol. 2, no. 9, pp. 338–347, 1998.

[7] C. Castellini, F. Orabona, G. Metta, and G. Sandini, "Internal models of reaching and grasping," *Advanced Robotics*, vol. 21, no. 13, pp. 1545–1564, 2007.

[8] D. Bullock, S. Grossberg, and F. H. Guenther, "A self-organizing neural model of motor equivalent reaching and tool use by a multijoint arm," *Cognitive Neuroscience*, vol. 5, no. 4, pp. 408–435, 1993.

[9] Y. Demiris and A. Dearden, "From motor babbling to hierarchical learning by imitation: A robot developmental pathway," in *Int. Conf. Epigenetic Robotics (EpiRob)*, 2005.

[10] A. D'Souza, S. Vijayakumar, and S. Schaal, "Learning inverse kinematics," in *IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2001.

[11] G. Sun and B. Scassellati, "A fast and efficient model for learning to reach," *International Journal of Humanoid Robotics (IJHR)*, vol. 2, no. 4, pp. 391–413, 2005.

[12] F. R. Reinhart and J. J. Steil, "Neural learning and dynamical selection of redundant solutions for inverse kinematic control," in *IEEE-RAS Int. Conf. Humanoid Robots*, 2011.

[13] M. Kawato, "Feedback-error-learning neural network for supervised motor learning," in *Advanced Neural Computers*, R. Eckmiller, Ed. Elsevier, 1990, pp. 365–372.

[14] M. I. Jordan and D. E. Rumelhart, "Forward models: supervised learning with a distal teacher," *Cognitive Science*, vol. 16, no. 3, 1992.

[15] J. Porrill, P. Dean, and J. V. Stone, "Recurrent cerebellar architecture solves the motor-error problem," *Proc Biol Sci*, vol. 271, no. 1541, pp. 789–796, 2004.

[16] M. Kuperstein, "Neural model of adaptive hand-eye coordination for single postures," *Science*, vol. 239, no. 4845, pp. 1308–1311, 1988.

[17] D. Caligiore, T. Ferrauto, D. Parisi, N. Accornero, M. Capozza, and G. Baldassarre, "Using motor babbling and hebb rules for modeling the development of reaching with obstacles and grasping," in *Int. Conf. Cognitive Systems (CogSys)*, 2008.

[18] J. Piaget, *The Origin of Intelligence in the Child*. Routledge and Kegan Paul, 1953.

[19] B. I. Bertenthal, "Origins and early development of perception, action, and representation," *Annual Reviews Psychology*, vol. 47, pp. 431–459, 1996.

[20] R. K. Clifton, B. A. Morrongiello, W. Kulig, and J. Dowd, "Developmental changes in auditory localization in infancy," in *Development of Perception, Vol. 1, Psychobiological Perspectives*, R. Aslin, J. Alberts, and M. Petersen, Eds. New York: Academic Press, 1981, pp. 141–160.

[21] H. Bloch and I. Carchon, "On the onset of eye-head coordination in infants," *Behavioral Brain Research*, vol. 49, no. 1, pp. 85–90, 1992.

[22] A. van der Meer, F. van der Weel, and D. Lee, "The functional significance of arm movements in neonates," *Science*, vol. 267, no. 5198, pp. 693–695, 1995.

[23] A. van der Meer, "Keeping the arm in the limelight: Advanced visual control of arm movements in neonates," *European Journal of Paediatric Neurology*, vol. 1, no. 4, pp. 103–108, 1997.

[24] C. von Hofsten, "Eye-hand coordination in the newborn," *Developmental Psychology*, vol. 18, no. 3, pp. 450–461, 1982.

[25] L. Rönnqvist and C. von Hofsten, "Neonatal finger and arm movements as determined by a social and an object context," *Early Development and Parenting*, vol. 3, no. 2, pp. 81–94, 1994.

[26] C. von Hofsten, "An action perspective on motor development," *Trends in Cognitive Science*, vol. 8, no. 6, pp. 266–272, 2004.

[27] R. K. Clifton, D. W. Muir, D. H. Ashmead, and M. G. Clarkson, "Is visually guided reaching in early infancy a myth?" *Child Development*, vol. 64, no. 4, pp. 1099–1110, 1993.

[28] L. Out, A. J. van Soest, G. P. Savelsbergh, and B. Hopkins, "The effect of posture on early reaching movements," *Journal of Motor Behavior*, vol. 30, no. 3, pp. 260–272, 1998.

[29] J. Baily, "Adaptation to prisms: Do proprioceptive changes mediate adapted behaviour with ballistic arm movements?" *The Quarterly Journal of Experimental Psychology*, vol. 24, no. 1, pp. 8–20, 1972.

[30] E. W. Bushnell, "The decline of visually guided reaching during infancy," *Infant Behavior and Development*, vol. 8, no. 2, pp. 139–155, 1985.

[31] J. Konczak, M. Borutta, H. Topka, and J. Dichgans, "The development of goal-directed reaching in infant: hand trajectory formation and joint torque control," *Experimental Brain Research*, vol. 106, no. 1, pp. 156–168, 1995.

[32] N. E. Berthier and R. Keen, "Development of reaching in infancy," *Experimental Brain Research*, vol. 169, no. 4, pp. 507–518, 2005.

[33] M. Rolf, J. J. Steil, and M. Gienger, "Goal babbling permits direct learning of inverse kinematics," *IEEE Trans. Autonomous Mental Development*, vol. 2, no. 3, 2010.

[34] E. Theodorou, J. Buchli, and S. Schaal, "Reinforcement learning of motor skills in high dimensions: A path integral approach," in *IEEE Int. Conf. Robotics and Automation (ICRA)*, 2010.

[35] A. Baranes and P.-Y. Oudeyer, "Maturationally-constrained competence-based intrinsically motivated learning," in *IEEE Int. Conf. Development and Learning (ICDL)*, 2010.

[36] W. Schenck, "Adaptive internal models for motor control and visual prediction," in *MPI Series in Biological Cybernetics*. Logos Verlag: Berlin, 2008, doctoral Thesis.

[37] C. G. Atkeson, "Learning arm kinematics and dynamics," *Annual Reviews Neuroscience*, vol. 12, pp. 157–183, 1989.

[38] E. Oyama and T. Tachi, "Goal-directed property of online direct inverse modeling," in *Int. Joint Conf. Neural Networks (IJCNN)*, 2000.

[39] T. D. Sanger, "Failure of motor learning for large initial errors," *Neural Computation*, vol. 16, no. 9, pp. 1873–1886, 2004.

[40] J. Peters and S. Schaal, "Reinforcement learning by reward-weighted regression for operational space control," in *Int. Conf. Machine Learning (ICML)*, 2007.

[41] M. Rolf and J. J. Steil, "Explorative learning of right inverse functions: theoretical implications of redundancy," in *New Challenges in Neural Computation*, 2012.

[42] M. Rolf, J. J. Steil, and M. Gienger, "Online goal babbling for rapid bootstrapping of inverse models in high dimensions," in *IEEE Int. Joint Conf. Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*, 2011.

[43] L. B. Smith and E. Thelen, "Development as a dynamic system," *Trends in Cognitive Sciences*, vol. 7, no. 8, pp. 343–348, 2003.

[44] U. Sailer, J. R. Flanagan, and R. S. Johansson, "Eye–hand coordination during learning of a novel visuomotor task," *Journal of Neuroscience*, vol. 25, no. 39, pp. 8833–8842, 2005.

[45] M. Rolf and J. J. Steil, "Efficient exploratory learning of inverse kinematics on a bionic elephant trunk," 2012, submitted.

[46] L. Jamone, L. Natale, K. Hashimoto, G. Sandini, and A. Takanishi, "Learning task space control through goal directed exploration," in *IEEE Int. Conf. Robotics and Biomimetics (ROBIO)*, 2011.

[47] P. O. Stalph and M. V. Butz, "Learning local linear jacobians for flexible and adaptive robot arm control," *Genetic Programming and Evolvable Machines*, vol. 13, no. 2, pp. 137–157, 2012.

[48] C. Hartmann, J. Boedecker, O. Obst, S. Ikemoto, and M. Asada, "Real-time inverse dynamics learning for musculoskeletal robots based on echo state gaussian process regression," in *Robotics: Science and Systems (RSS)*, 2012.