The goal of this proposal is to make progress on computational problems that elude the most sophisticated AI approaches but that infants solve seamlessly during their first year of life. To this end we will develop an integrated system (a child robot) whose sensors and actuators approximate the levels of complexity of human infants. The target is for this robot to learn and develop autonomously a key set of sensory-motor and communicative skills typical of 1 year old infants. The project will be grounded in developmental research with human infants, using motion capture and computer vision technology to characterize the statistics of early physical and social interaction. An important aspect of the proposal is the effort to integrate developmental processes from very different domains under a common framework. The proposed projects explore how principles from control theory can drive the robot to learn how to control its body, to learn about physical objects, and to learn how to interact with people.

**Scientific Impact:** Artificial intelligence, machine learning, and machine perception have made good progress by focusing on specific tasks, e.g., chess playing, face detection, expression recognition, character recognition, speech recognition, unsupervised object discovery, document analysis [7, 50, 51, 53, 87, 107, 113]. As performance on these component tasks improves, the need to understand how to bring them together into a more broadly competent system has become apparent. Roboticists and cognitive scientists sometimes refer to this issue as the “architecture problem”, which was first emphasized by Allen Newel [81]. A variety of cognitive and behavioral architectures have been proposed [3, 4, 12, 46, 48, 82] however they are typically static and need to be preprogrammed by hand. Autonomous learning and development plays a minimal, if any role. By now, it is becoming clear that these approaches may not scale up as the complexity of sensors and actuators approximates that of human beings [115]. Conceptual shifts are needed to rigorously think, explore, and formalize architectures that learn and develop autonomously by interaction with the physical and social worlds. The goal of this proposal is to start making progress towards this end. The approach we propose is to reverse engineer the developmental process human infants go through during the first year using a unique dataset that traces the development of key sensory-motor and social behaviors during the first year of life. In order to better understand the computational problems faced by infants and the developmental solutions they find, we will focus on the task of controlling a state of the art humanoid child robot (CB2) with rich human-like sensors and actuators (See Figure 1). The complexity of the robot defies traditional approaches and presents unique opportunities to explore solutions that rely on learning and developmental processes.

The research team consists of: (1) **Javier R. Movellan (Lead PI)**, Research Professor. Institute for Neural Computation, University of California San Diego (UCSD). Research interests focus on machine learning, machine perception, robotics, computational models of learning and development. (2) **Daniel Messinger (PI Miami Team)**, Associate Professor, Department of Psychology an Pediatrics, University of Miami. He is a prominent developmental psychologist focused on understanding the development of real-time physical and social interaction during the first year of life. (3) **Emmanuel Todorov (Co PI)** Assistant Professor Cognitive Science. UCSD. Research interests focus on biological motor control, machine learning, and computational neuroscience. (4) **Virginia de Sa (Co PI)** Assistant Professor, Department of Cognitive Science, UCSD. Research interests focus on machine learning, computational neuroscience and multi-sensory information integration. (5) **Marian S. Bartlett (Co PI)** Research Scientist. Institute for Neural Computation. University of California, San Diego. Research interests focus on computer vision and its application to the analysis of social behavior. (6) **Hiroshi Ishiguro (Consultant)** Professor. Intelligent Robotics Laboratory. University of Osaka. Research interests focus on humanoid and android robots. (7) **John Watson (Consultant)** Professor Emeritus. Department of Psychology. University of California Berkeley. Learning in infancy. (8) **Terrence Sejnowski (Consultant)** Professor, the Salk Institute. Computational Neuroscience. We are a unique interdisciplinary research team, composed of experts in robotics (Todorov, Ishiguro, Movellan), machine learning (de Sa, Bartlett, Todorov, Movellan), biological motor control (Todorov), computer vision (Bartlett, Movellan), and infant development (Messinger, Movellan). The lead
researcher, Javier Movellan, is ideally positioned to foster the interdisciplinary interactions that are needed to make this project successful. He has has a Ph.D. in developmental psychology and his research work spans machine learning theory [58, 64, 65, 68, 72, 74, 75, 79], machine perception [5, 6, 7, 8, 9, 29, 38, 41, 54, 59, 60, 66, 69, 71, 72, 77, 108], robotics [14, 27, 28, 32, 47, 55, 56, 67, 78, 97, 98, 99], and behavioral science [28, 70, 76, 93].

Figure 1: The CB2 humanoid robot (Child-robot with Biomimetic Body). It has 56 highly compliant pneumatic actuators and a whole-body soft skin with an embedded network of tactile sensors. High resolution cameras are mounted inside the eyeballs. Microphones are mounted on the ears.

The proposed research will proceed in close collaboration with a twin project, led by Dr. Minoru Asada and Dr. Hiroshi Ishiguro at the University of Osaka, Japan. This project is being funded by the Exploratory Research for Advance Technology initiative of the Japanese Science and Technology Agency. The Japanese team will build a replica of the CB2 child robot currently used on their project. The new robot will be hosted by the California Institute of Telecommunications and Information Technology at UCSD. The collaboration between the Japanese and the US team will include internships for US students to work in Japan and for Japanese students to work in the US, joint scientific publications, joint data and software repositories, and joint workshops to discuss progress, identify challenges and set joint goals.

**Broader Impact:** The project will lay the foundation for an area of research critical for machines to achieve broader and more significant levels of intelligence that elude current approaches. Progress in this area could result, for example, in robots that assist people in daily life and interact with them in a personal manner, automatic tutors that understand human behavior and adapt to the needs of their students accordingly, and industrial robots of levels of complexity unachievable with current approaches. The project may also open new avenues to the computational study of infant development and potentially offer new clues for the understanding of developmental disorders such as autism and Williams syndrome. This work will also help to forge stronger links between behavioral scientists studying human learning and development and computational scientists studying machine perception and learning. Solid links between these areas have the potential to greatly improve our knowledge of both areas.

The project has a strong outreach and educational component. It will fund 6 graduate students, 1 postdoctoral student, and 2 junior faculty members. As part of an ongoing collaboration, we will be involved and collaborate with the Robotics Team and the Engineering Course at the Preuss School in San Diego. This is a charter School for low-income student in grades 6-12. The demographics at Preuss
2004/05 were: 59.5% Hispanic, 21.7% Asian, 12.9% African American, 6% White. Last year more than 95% of their graduating seniors moved on to 4 year Colleges around the country. The students will have internships at UCSD to work with the research team. Efforts will be made to invite K-12 Schools across the nation to interact with the researchers and the robot (physically and via Web) as the project progresses.

1 Description of The Overall Mathematical Framework

![Diagram](image)

Figure 2: The Bayesian Decision Theory/Optimal Control view on motor control. The utility/cost function defines what the person is trying to achieve, including how valuable it is to make fast and accurate movements and how costly it is to expend energy. A Bayesian subsystem infers the state of the world from sensor readings. The controller combines the probabilistic information about the world with the cost function to compute the best motor commands.

Formally our project is grounded on Bayesian inference and the theory of stochastic optimal control [102, 104], particularly contemporary machine-learning based approaches that have been successfully applied to complex control problems [92, 100, 101, 103, 105, 109, 110]. The brain faces a control problem when sending motor commands to the limbs. In order to move the hand to a desired position in space, it needs to take into account the inertial properties of the arm and levels of uncertainty in the efferent and afferent transmission lines. Using a computer mouse, riding a bicycle, shooting baskets, and playing a musical instrument are also control problems. In addition to the properties of the limb, the brain has to take into consideration the control properties of the tool it interfaces with. We propose that learning to interact with others in real time is in essence also a control problem and that the mechanisms responsible for learning to control our bodies and to control physical objects are also responsible for learning to interact with others.

In essence the control theory approach postulates that knowledge comes down to anticipating what we will sense as a function of what we do. Representations are compressed versions of sensory-motor history that are useful to help predict the world. When an infant displays ritualized reaching behavior towards an object expecting that a caregiver will hand it down he is displaying very sophisticated knowledge of the control properties of human beings. Understanding how this knowledge may be acquired in an autonomous manner is a key objective in this project. Defining knowledge in terms of relationships between sensors and actuators grounds it in the physical world and enables us to use the powerful machinery of modern machine learning, and stochastic optimal control.

Formally control theory is a discipline of mathematics and engineering that deals with the control of probabilistic dynamical systems. In a typical stochastic control problem we are given a dynamical
system, commonly as the world or the plant, that we need to control:

$$X_{t+\delta t} = X_t + \mathcal{M}(X_t, U_t, W_t)$$  \hspace{1cm} (1)$$

where $X_t$ is the state of the system at time $t$, $U_t$ is a control signal that the controller can use to “steer” the state, and $W_t$ is a noise process that makes the system stochastic. The function $\mathcal{M}$ known as the system function determines the state dynamics. Depending on the situation it is may be convenient to model the control problem as a discrete time or a continuous time process. In discrete time processes, $\delta t$ is a positive number. Continuous time processes are typically modeled using stochastic differential equations, which can be seen as a limiting process as $\delta t \rightarrow 0$.

The complexity of control problems can be measured in terms of the degree of knowledge available to the controller. In the simplest case the controller has direct access to the state of the world $X_t$. These are known as Markov Decision Processes (MDP). A more realistic, and more difficult problem, occurs when the controller knows the world dynamics but needs to infer the current state of the world. These types of control problems are known as Partially Observable Markov Decision Processes (POMDP). In such problems the controller has access to sensor measurements that are controlled by a known sensor function $S$:

$$Z_t = S(X_t, W_t)$$  \hspace{1cm} (2)$$

and has to use the history $H_t$ of sensory measurements and control signals to estimate $X_t$ and to achieve the desired goals. For example, $H_t$ could contain the images, sounds, tactile, and inertial information available to the robot, and $X_t$ could encode the robot’s map of the world, including its own location in this map and the location of objects of interest.

Finally, the most difficult type of control problem occurs when the system needs to infer the state of the world $X_t$ and the world dynamics $(\mathcal{M}, S)$. This can be elegantly modeled within a Bayesian framework by treating $\mathcal{M}$ and $S$ as random variables that need to be inferred from sensory-motor data. Such problems are known as Bayes-Adaptive POMDPs.

System's identification refers to the problem of making inferences about $\mathcal{M}$ and $S$ based on the sensory motor history $H_t$. System’s control refers to the problem of achieving goals within this world.

Mathematically a controller (aka a policy or a control law) is a function $\pi$ that maps moment-to-moment the information history $H_t$ into actions $U_t = \pi(H_t)$. An optimal controller $\hat{\pi}$ optimizes a performance function $\rho$ expressed as the long term accumulation of a reward or utility function $r$

$$\rho(\pi) = E \left[ \int_0^T r(X_t, U_t) dt \mid \pi \right]$$  \hspace{1cm} (3)$$

where $E$ is the expected value operator, and $T$ is the, possibly random and possibly infinite operation time. For example, $X_t$ may represent the location of a robot arm and the goal of the controller may be to reach a desired target $\omega$. In such case the utility may decrease proportionally to the Euclidean distance from the target and to the energy expenditure. Algorithms for finding optimal controllers exist for a few special cases, e.g., when the world dynamics are linear the reward function is quadratic. However for most problems of interest in robotics and motor control, exact solutions are too computationally expensive and most of the work in this area of research has focused on approaches to find approximate solutions. These approaches include neural networks, reinforcement learning, and differential dynamic programming [92, 100, 101, 103, 105, 109, 110]. The last few years have seen great progress towards the development of practical algorithms for POMDP and BAPOMDPs [84, 85, 89]. These algorithms have emerged from recent interaction between the robotics and machine learning communities and can handle problems large enough to tackle real life robotic applications. Emanuel Todorov, a co-PI in this project, is a leader in the application of the theory of stochastic optimal control to problems in biological motor control [104]. To this effect he has developed novel algorithms for approximate solutions to control problems using realistic muscle models with large numbers and a large number of degrees of freedom [100, 101, 103, 105, 109, 110]. Javier Movellan, the lead PI is a pioneer in the use of control theory framework to understand the development of early social interaction [67, 73].
2 Previous Work: RUBI

The proposed activities were in great part motivated by and will build upon our experience on the RUBI project (http://mplab.ucsd.edu). The goal of this project is to explore the idea of robots that teach children skills and assess their development while interacting with them in an affective and human-like manner. The main philosophy in this project is to design robots by immersion in daily life environments. To this effect for the last 3 years we have been immersing robots, and researchers at an Early Childhood Education Center at the University of California San Diego. Results from this project were published at major scientific journals, including the Proceedings of the National Academy of Sciences (in press), received a best paper award at the IEEE International Conference on Robot and Human Interactive Communication, and led to smile detection that has been ported to a new generation of commercial digital cameras.

Many of the perceptual primitives for social robots developed as part of the RUBI project will be incorporated in the new proposed robot, these will include automatic face detection, facial expression analysis, social contingency detection, and auditory mood recognition [7, 67, 91]. In addition, we have developed a software architecture for social robots, named RUBIOS, that is based on the principles of Bayesian inference and stochastic optimal control [27]. Each node of a RUBIOS robot keeps track of a model of the world, and an internal measure of success. RUBIOS nodes can interact with other nodes by means of temporal bids, i.e., promises to provide “utility” to other nodes provided particular goals are accomplished at specified points in time. RUBIOS is on its second revision now and it provides the basic software architecture controlling the robots in the RUBI project. RUBIOS will be ported to the new proposed robot.

Most importantly, our experience developing robots that have to interact with children autonomously for extended periods of time, has helped us realize the need for fundamental research on robots that develop on their own by interaction with the world. This is the main focus of this proposal. The RUBI project is synergistic with the research project proposed. Th RUBI project helps us realize the problems that need to be solved for daily life robots to become a reality, the current project focuses on the fundamental research that needs to be done to solve such problems.

Figure 3: Each year of the RUBI project we developed a different robot prototype that incorporated the lessons learned from the previous years. Left: RUBI-1 was partially controlled by a human operator. Center: RUBI-2 goes for a nature walk with the children. Right: RUBI-3 Playing educational games. It operates fully autonomously for weeks at a time.

3 Previous Work: CERT

Marian Bartlett, a Co-PI in this project has developed a system, named CERT, for automatic analysis of facial expressions from video. The system automatically detects frontal faces in the video stream and codes each frame with respect to 37 continuous dimensions, including basic expressions of anger.
disgust, fear, joy, sadness, surprise, as well as 30 facial action units (AUs) from the Facial Action Coding System [26]. CERT was trained on multi-terabyte databases of spontaneous and posed facial expressions and achieves unmatched performance in real-time at video frame rates [7]. Recently, CERT was successfully employed to discriminate facial expressions of real pain from posed pain, and to learn new associations between facial behaviors and driver fatigue [108]. Most importantly CERT provides information on the dynamics of facial expression at high spatial (30 dimensions of expression) and temporal resolution (30 Hz).

CERT will play an important role in this project. First it will become one of the perceptual primitives for the proposed robot, thus giving it access to detailed information about the facial expressions of human caregivers. Second, it will be used to analyze the development of facial expression dynamics between infants and caregivers during the first year of life (See Section 7.1).

**Figure 4:** Left: Example smiles automatically extracted by CERT from the camera of the RUBI-3 robot. Right: Some of the complete images as viewed by RUBI-3.

### 4 Previous Work: SBFs

We recently introduced a general framework, named Segmental Boltzmann Fields (SBFs), for learning to detect and localize objects in natural scenes [30]. The framework allows the automatic discovery of the visual appearance of object categories from images that have been labeled only as containing, or not containing the object of interest, without indicating where the object is on the image. The approach generalizes and overcomes limitations of previous approaches in the computer vision literature: Markov Random Fields [10, 15, 35] and Fields of Experts [42, 90, 114]. SBFs model images as a mosaic of image segments, each of which is independently rendered by a different object. This provides an abstract yet powerful definition of object: a collection of random variables (e.g., a group of pixels) that are mutually dependent while independent of the other observed variables (e.g., the rest of the pixels in an image).

While the SBF framework was developed from the point of view of Bayesian decision theory we showed that resulting inference algorithms can be efficiently implemented as a three hidden layer convolutional neural network [52, 77]. For a 640 × 480 pixel retina this results on a network with a total of 5 million simulated neurons. Due to careful optimizations, the network runs comfortably in real time at 30 frames per second in a standard laptop, making it ideal for robotic implementations.

Given a set of images labeled as containing or not containing an object of interest, the goal in SBFs is to learn a model that allows to identify the object’s presence, as well as its location, in new images. Under the model learning consists of adjusting the network synapses so as to “make sense” of the
observed images. In probability theory this translates into maximizing the joint log probability of a
collection of observed images and its labels. Modern machine learning approaches, like AdaBoost
[33, 34] are used to perform this maximization very efficiently.
Perception in SBFs consists of making inferences about the observed images, given the network
parameters. The output of this inference consists of a posterior probability field, that describes the
probability of each pixel rendering the objects of interest. Due to the structure of SBFs, it is possible to
perform this inference in real time at little computational cost.
SBFs were tested on two datasets that have become standard in the computer vision community: Caltech
4 and Caltech 101 [31]. The first requires learning 4 different object categories, and the second 101
object categories. The SBF approach produced state-of-the-art accuracy results but was orders of
magnitude faster than the previously published approaches [31, 94], allowing real time operation at
standard video frame rates. Figure 5 shows examples of the inferences made by an SBF that learned to
detect leopards.

![Figure 5: Examples of inference on novel images in the "Leopards" category. The image above shows
the best classification window inferred by the model. On the bottom is the saliency map, showing the
approximate posterior probability under the model that each pixel was generated by the foreground
object.]

Unsupervised Discovery of Caregivers  Arguably the most critical object category for infants to
discover is that of human caregivers. There is strong experimental evidence that by 40 minutes of age,
neonates already orient towards human faces [37, 63], and that by 2 days of age they fixate longer to
images of their mothers than to images of other women with similar hair colors and facial complexion
[13]. In order to investigate how this learning process may occur we performed an exploratory
experiment with a simple baby robot built at our laboratory with off-the-shelf parts. The baby robot was
endowed with an SBF network consisting of 5 Million neurons. The robot vocalizations were
scheduled moment-to-moment by an optimal controller designed to infer as quickly as possible whether
a responsive human being was present [80]. Based on the history of auditory information the controller
periodically vocalized and when a auditory contingency was detected, the current retinal image was
automatically sent to the SBF learning module with the label contingency present. Otherwise the image
was sent with the label contingency absent. The robot’s task was to discover the visual appearance of
the objects that caused auditory contingencies.
After less than 6 minutes of interaction with members of the laboratory, the robot achieved a
performance level of 86.17% on face detection tasks in novel images and and 92.3 % correct on person
detection tasks in novel images. A remarkable aspect of this result is that during the 6 minutes of
exposure to the world, the baby robot was never told whether or not people were present in the images,
or whether people were of any particular relevance at all. The robot could actually generalize to
abstract face image stimuli that had been used in classic experiments on neonate face perception [45].
Figure 6: Examples images and their pixel-wise probability images, indicating the posterior probability that each pixel was generated by an object. If the likelihood ratio of an image containing object vs. no-object exceeds a threshold, the best classification window is drawn. On the top row are good localization and detection results, despite variations in lighting, scale, gender, pose, and facial expression. Note that the top right image is an example that was originally labeled “not contingent”. From left to right on the bottom row: (1) correct rejection, (2)-(4) correct detections, where the body was preferred over the face, (5) the most probable location was incorrect, however the image was correctly classified, (6) an incorrect rejection, (7)-(8) incorrect detections.

This results was highlighted in the front-cover of the APA Monitor for its potential to transform the scientific understanding of infant development.

The experiment also provided an important lesson for us: We have experience developing computer vision for automatic face detection and expression recognition [5, 6, 7, 8, 9, 29, 54, 59, 60, 77, 108]. Training these systems typically requires tens of thousands of images of faces carefully segmented and labeled by hand. Here we were asking a robot to autonomously solve what appeared to be a much harder problem: to discover how humans look based on simple auditory contingencies, without any segmentation labels. Yet the problem was solved faster than we had anticipated, illustrating that complex problems are not always harder if one approaches them the right way. As it turned out key to the robot success is that unsegmented images that have faces in them have much more information than the segmented patches of faces we had used in previous approaches. This is a key motivating theme for this proposal: The belief that breakthroughs can happen by approaching very difficult perceptual and control problems in a manner that reverse engineers the solutions infant seamlessly find via developmental processes.

5 Proposed Research: Active Object Discovery

When infants are born, very few of their perceptual skills are fully developed. However, after two days of life they already respond differentially to faces of their caregivers [13] and within one year they can recognize a large number of objects and interact with them in many different situations [83]. Understanding how this is possible and developing machines that can develop autonomously in a similar manner could have profound scientific and technological implications. We have experience working on this area of research and believe significant progress is possible within the next four years. While our previous work in this area illustrates the type of studies we would like to pursue, it had many limitations: The robot has a single actuator (a microphone) and a single sensor (a camera). It could not actively move to touch objects or to orient its camera in the direction of interesting stimuli. As part of this project we propose to address these limitations towards the goal of demonstrating a robot (CB2).
that can learn without any supervision hundreds of object categories using multiple modalities. The actuators will include vocalizations, arm/hand movements, and head/eye movements. The sensory modalities will include audition, vision, touch, and inertial sensors. We propose to progress towards this goal through a series of studies.

**Study 1:** This study will focus on the discovery of object categories using hand and arm movements, and sensory information from audio, video, touch and proprioception. The robot will grasp objects using built in reflexes and then will select, using a prebuilt control policy, a sequence of behaviors typically observed in infants (e.g., shake object, take object to mouth). Five different categories of objects will be chosen from those likely to be encountered by a typical infant (e.g., a water bottle, a piece of cloth, a teether, a rattle, the hand of a human caregiver). The touch information, auditory signals, and proprioceptive information will be recorded and online non-parametric Bayesian clustering methods will be applied to these signals [1, 39, 88, 116] . The advantage of non-parametric Bayesian methods over classical clustering approaches like k-means, are that: (1) they do not require a predefined number of clusters, instead they make inferences about the most probable number of clusters as well as the structure of these clusters; (2) They operate incrementally rather than in batch mode (i.e., inferences change continuously as new sensory data arrive). CB2 will be programmed to move grasped objects so they are visible by its cameras. The cluster labels obtained using the non-visual modalities will then be used to train the robot’s visual system (an SBF network). For example, based on auditory, tactile and proprioceptive information an object may be recognized as being part of Cluster A.

One disadvantage of SBFs is that they need an external system to inform them whether or not an object of interest is present. Virginia de Sa, a co-PI on this grant, has worked on alternative algorithms that allow different sensory modalities to train each other in an unsupervised way [17, 18, 19, 20, 21, 22, 23, 24]. The algorithms use the correlations between different sensory modalities to steer the learning process. As different modalities are sensitive to different environmental changes, they can be useful to help each other discover invariant features for recognition. The idea is to adjust the category boundaries in each modality to minimize the disagreement between the resulting categories output by each modality. We will compare the above SBF method with de Sa’s Minimizing-Disagreement (M-D) algorithm using the same data. After training with each of the methods the robot will be tested to evaluate whether it can recognize visually the objects it interacted with. To do so the robot will need to have learned a pose independent representation of each object and to filter its own hand out of this representation (i.e. during training the images of the objects include the robot’s hand that grasp the object, during testing we will test the object images without the hands). We will then analyze the advantages and disadvantages of each method and, if indicated, attempt to develop a hybrid algorithm that has sophisticated SBF classifier but uses information in all directions like the M-D algorithm.

**Study 2:** In study 1 the robot was endowed with reflex behaviors designed to maximize the discriminability of available objects (e.g., grasp objects that touch the hand, shake the objects so they produce auditory information, move the hands so the objects are visible). In study 2, we will investigate whether these behaviors can also be acquired in an autonomous manner. Formally the goal is to develop a controller for the task of multi-modal object discrimination. We already have experience developing such controllers from examples using standard reinforcement learning approaches like TD-Lambda [44, 95]. In this case the reinforcement signal will be the amount of information provided by a sequence of behaviors from the arms and hands, where information is defined with respect to the learned object clusters. For example, if all the objects happen to be rattles of different colors but otherwise identical, shaking the rattle shall not provide any discriminative information, but moving the object close to the cameras will. The amount of discriminative information provided by a behavior can be easily measured and used as a reinforcement signal to provide a controller that maximally discriminates the available objects [44].
Study 3: In study 3 we will extend the previous studies using a realistic number of objects, a realistic distribution of objects, and time scale of several weeks. During the study CB2 will be given the opportunity to interact with hundreds of objects chosen to represent a typical infant’s life. This will include both physical, and social objects (i.e., undergraduate students hired to interact with the robot on a daily basis). During the study we will track the representations learned by the robot as learning progresses (e.g., the way it clusters objects as learning progresses) as well as the behaviors it develops so as to optimally discriminate the objects in its life.

6 Proposed Research: Sensory Motor Control

Different motor behaviors are often studied in isolation, despite the fact that they normally occur in a coordinated fashion and are likely mediated by overlapping neural mechanisms. Examples of such research niches include eye movements, reaching, grasping, walking, postural maintenance. It may be possible to build a rich behavioral repertoire by first learning isolated sensorimotor skills and later putting them together. However that does not correspond to the pattern observed during development. Instead, even the earliest attempts to reach and grasp are accompanied by eye, head and body movements directed to the same spatial goal (an object of interest). Later in life these coordination patterns continue to be refined. Here we propose to investigate sensorimotor learning from a single unifying perspective - stochastic optimal control - with the expectation that the same task will give rise to multiple behaviors in different contexts. A motivating example is the work of Karl Sims who used genetic algorithms to optimize various simulated creatures for the task of swimming. The striking outcome of these simulations was the emergence of life-like diversity of behaviors pursuing a common goal (propel the center of mass forward). Every arm, leg, tentacle and other hard-to-name body part of the simulated creature ended up moving in such a way that the entire creature made progress towards the goal, in apparent disregard of the behavioral partitioning which the field of Motor Control has adopted. The spirit of the work proposed here is similar, with the following differences: (i) we will focus on human movement; (ii) we will use not only realistic simulation but also an advanced robotic platform; (iii) we will use more powerful learning algorithms which exploit the structure of the underlying problem; (iv) we will use insights and experimental data from developmental psychology to guide the computational work.

Learning to Reach: Our goal will be to learn a control law (a mapping from body states to actuator commands) which makes the robot achieve a stable grasp around an object. Task achievement will be encoded by a suitable cost function. By presenting objects of different shapes, at different positions and orientations relative to the body, while the body is in different initial configurations, we will elicit a rich set of behaviors. If the object is within arm’s reach a coordinated arm and hand movement will be sufficient, although even in this case one has to compensate for interaction forces throughout the body. If the body is lying on the ground and the object is presented sufficiently high, grasping it will require
transition to a sitting posture before an arm movement can be successful. Similarly, if the object is to the side, the body may need to be turned and reoriented. An even higher elevation of the object will require standing and perhaps jumping. If the object is across the room, grasping it will require crawling or walking or some other form of locomotion. At all times, orienting head and eye movements will be necessary in order to keep the object in view. Thus a seemingly simple task can elicit a wide range of motor behaviors by varying the spatial parameters of the task. Importantly, these task variations correspond to experiments that can be done with infants at different ages. Such experiments will be designed and performed here, and will generate a unique database of infant movements in the context of a single yet rich task. This data will then guide the development of learning algorithms for the robot. In particular, we will analyze quantitatively how the motor strategies adopted by infants evolve with age as well as experience in the experimental protocol. This analysis should reveal a progression of stages, which will then be used to fine-tune our learning algorithms so as to mimic human development and consequently (we believe) result in better task performance.

We know well that the control problems we propose to study are extremely challenging. None of the analytic techniques in classical control theory are likely to work here. Instead we will harness the power of modern computers, and use numerical optimization to learn controllers that are more intelligent than anything we can build by hand. A variety of such methods are available in related fields: Reinforcement Learning, approximate dynamic programming, neuro-dynamic programming. These methods start with a guess about the control policy or the value function (the function which indicates what long-term performance can be expected if the system is initialized in a given state). This guess is then iteratively improved. Grid-based methods suffer from the "curse of dimensionality" - which refers to the fact that the volume of state space grows exponentially with dimensionality. Instead of grid-based methods we will use either function approximation methods, or local trajectory-based methods whose complexity scales quadratically rather than exponentially. We have extensive experience with both [100, 101, 103, 104, 105]. A particularly promising class of approximations to optimal control is hierarchical feedback control [110]. The idea is to construct a low-level feedback controller which performs a transformation of the world dynamics, resulting in an augmented world model that is easier to control. A high-level controller can then focus on solving the task in a reduced space. We and others have developed such methods and found them to be very efficient in practice. An example is shown in Figure 7 where we used our hierarchical control method to compute the activations of 20 muscles acting around a detailed human arm model. The task was to move the arm up to an elevated position and then bring it down to an intermediate position. Hierarchical control is appealing not only because it simplifies complex problems, but also because there is extensive experimental evidence that the sensorimotor system relies on motor primitives, or motor synergies, or control hierarchies, or some such mechanism. There is little agreement as to the exact nature of this mechanism. However most researchers in Motor Control agree that a mechanism with the above intuitive properties exists in the brain and is key to generating complex behaviors. Here we will experiment with different candidates for hierarchical control and compare them on the basis of the resulting performance as well as their similarity to the experimental data. We will also apply dimensionality reduction techniques to the data and attempt to identify the motor synergies used by infants.

Optimizing a control policy requires extensive trial-and-error which cannot be done with the real robot. Therefore we will develop an accurate dynamics simulator, and fine-tune it by sending different control signals to the robot, measuring the resulting movements and applying system id. We have experience developing simulators for musculo-skeletal systems such as the one illustrated in Figure 7. This simulator computes the 3D muscle geometry using virtual obstacles (blue objects) to constrain the muscle paths. It then computes the articulated-body dynamics with an efficient algorithm. Our experience with developing such tools, as well as the availability of reusable code, will allow us to build a dynamics simulator which can capture the behavior of the robot. This will enable learning of control policies in simulation. Then the resulting controller will be applied to the real robot and further adapted. The necessary adaptation will require much less trial-and-error than the initial learning, so it will be feasible to do it on the real robot.
7 Proposed Research: Social Interaction

The first year of life involves many challenges. Motorically, infants progress from uncoordinated head movements to the onset of walking. Cognitively, infants progress from relatively simple matching of body movements and afferent sensation to the beginnings of representation of absent objects. Socially, infants progress from recognizing the faces of others to intentionally communicating their desires through gestures.

The working hypothesis in this project is that the same principles that govern how we learn to control our own bodies also govern how we learn to interact with others. Both can be cast as problems in systems identification and real-time control (See Figure 8). This view was inspired by earlier work by John Watson, a consultant for this project. He proposed that early in development infants actively identify objects and humans by the way they react to the infant, i.e., by their control properties [112]. This approach originated in an experiment in which 2-month-old infants learned to move their heads on a pressure sensitive pillow to activate a mobile above their cribs. After 4 days of exposure to this controllable mobile, infants exhibited social smiles, positive affect, and cooing when the mobile was present. These social behaviors appeared notably less in a control group for which the mobile moved in a non-contingent manner. Watson proposed that the infants were basically treating this mobile as a proto social object.

Figure 8: Three real time control problems: person controlling a computer mouse, infant playing smile games with Mom, a child directing the line of regard of a robot.

Since then evidence has been found to support the idea that infants have very sophisticated knowledge of the control properties of other humans [86, 111]. For example by 5 months of age infants are tuned to the levels of responsiveness found in the interaction with their specific caregivers [11]. By the end of the first year, infants exhibit a variety of behaviors that reflect a rather sophisticated understanding of the control properties of other human beings. Indeed infant behavior with objects and others shows predictable patterns characterized by timing (e.g., latency of gaze shifts), sequence (e.g., puzzled to smiling), and content (e.g., preference for faces and motion). Infants from 2 to 12 months simultaneously produce and learn these regularities, which we believe are key to help understand typical and atypical social development. Our goal in this project is to begin to understand how such sophisticated social knowledge could be learned from sensory-motor experience in an autonomous manner. We believe understanding this process is critical to develop machines that can interact with people naturally in everyday life. To pursue this goal we propose the following activities: (1) Collection and datamining of rich datasets of infant/caregiver interaction. (2) Development of realistic perceptual and motor primitives for social robots. (3) Development and evaluation of algorithms for complex robots to learn to interact with humans on their own.

7.1 Parent Infant Interaction Dataset ($P I^2$)

Progress in the computational understanding of infant social development requires sensor rich datasets that capture the dynamics of early social and physical interaction. Previous work by Daniel Messinger, the PI of the Miami team, has already shown the value that such datasets could bring to the computational study of early social interaction [117]. Here we propose to extend this work and to collect...
an unprecedented dataset of Parent Infant Interaction (the \textit{PI}^2 dataset) using state of the art motion capture and computer vision technology. The \textit{PI}^2 dataset will be collected at the proposed Infant Motion Capture Laboratory at the University of Miami. Movellan, Todorov, and Bartlett, have already set up a similar facility at UCSD, including development of software for robust capture of the human body [103] and for synchronization of motion capture and video based analysis of facial expressions. They will assist the Miami team to setup the proposed facility in Miami.

The proposed dataset will contain synchronized motion capture data (at 240 Hz), and facial expression analysis from video (at 30Hz) simultaneously obtained for infant/caregiver dyads. The motion capture will focus on the position and orientation of the head, arms, and hands. The CERT expression recognition system will be used to code dyadic interactions with near frontal camera views (less than 15 degrees rotation in depth). When necessary, manual coding of facial expressions will be performed by a certified FACS coder (Katherine Kern). The dataset will also include computer assisted coding of vocalizations, haptic behaviors (e.g., when mother touches baby) and gaze direction. The Miami team has already developed software tools to facilitate such coding.

The dataset will consist of 1650 minutes of early social interaction. This will include 10 caregiver/infant dyads (typically developing) studied longitudinally at one month intervals between 1 and 12 months of age. Parents will engage in the following types of interaction with their infant for 5 minutes each: (1) Face-to-face interaction; (2) Peekaboo; (3) Reaching to objects in the presence of a caregiver. Dr. Messinger, the PI of the Miami team, has extensive experience conducting longitudinal studies of this type [25, 43, 61, 62, 117]. The proposed dataset will be, to our knowledge unprecedented in that it will combine motion capture and automatic analysis of facial expression. In addition to its value for the project, the dataset is likely to become an important tool for the study of human development.

As the project progresses we will conduct supplemental cross-sectional studies of infants. These studies will focus on gathering additional data and testing hypothesis generated in the process of developing the robot model. These additional studies are expected to involve more frequent observations of infants over periods of 2-3 months that may focus on a specific developmental process (e.g., the development of socially mediated reaching).

\subsection{7.2 Modeling Social Expressions}

First we will focus on the analysis of the facial expression dynamics infants exhibit throughout development. This will include frequency counts of the most commonly used Facial Action Units (AUs), but most importantly a description of the frame-by-frame dynamics of these units. Such frame by frame analysis is unheard of and we can only make use of it thanks to the use of computer-vision that we have recently develop.

These data will be used to improve the facial expression dynamics of the proposed robot, with our focus on creating a continuum from positive expressions to negative expressions. In infants this continuum is known to involve the following units from the Facial Action Coding System: AU4 (brow furrow), AU12 (smile), AU6 (eye wince), and AU26/27 (jaw drop and mouth stretch). The US team will work with the Japanese team to develop actuators to better simulate the muscles involved in these AUs and to produce expression dynamics similar to the ones observed in the \textit{PI}^2 dataset.

We will also explore the joint dynamics between AUs and different modalities. To this end Virginia de Sa will use the Minimizing-Disagreement Algorithm she developed in collaboration with Dana Ballard [17, 23]. The goal will be to find clusters of gestures, vocalizations and facial expressions that are mutually predictive. Results will be compared across development and across the 3 different situations targeted in the PI2 dataset.

\subsection{7.3 Inverse Optimal Control}

The second and perhaps most critical function of the dataset is the extraction of models of early social interaction that can be used to generate and evaluate optimal control models for the proposed robot.
This will require two parallel efforts: (1) Reverse engineer the cost/utility function optimized by infants during social interaction episodes; (2) Reverse engineering (systems identification) of the response properties of caregivers in early social interaction.

**Reverse Engineering the Cost/Reward Function:** Even in simple situations, like reaching to a desired location in space, the actual function optimized by humans is subject to intense debate [40, 57, 96, 104]. There are motor costs, such as the metabolic energy used by a forceful movement as well as the resulting fatigue. There are also task costs, such as the time wasted in a very slow movement or the undesirable consequences of failing to properly grasp a glass of water. There are information costs, like the fact that moving an arm in front of the eyes may block important information. The situation becomes even more complex in social interaction. In previous work we showed that the vocalizations made by 10 month infants during the first minute of interaction with a new social agent maximize the information gained about the responsiveness of the agent [67]. While information gain may play an important role in many other social situations, it is likely the case that infants optimize a complex composite of costs and rewards. Here we propose to use the $PI^2$ dataset to reverse engineer the goals infants have when interacting with their caregivers, and then build robots that learn to achieve those goals on their own.

The problem of inferring the cost/utility function optimized by a controller from its observed behavior is known as inverse optimal control. One of the most elegant approaches to inverse optimal control was recently proposed by Konrad Kording and Daniel Wolpert [49]. Emo Todorov, a Co-PI for this project is collaborating with Kording, with whom he as developed an improved (yet unpublished) algorithm. Andrew Ng’s group has proposed an alternative method, which they named Inverse Reinforcement Learning particularly well suited for discrete state problems [2]. Their approach was successfully used to infer the function optimized by humans while driving a simulated automobile. Here we propose to apply these methods to infer the reward function optimized by infants at different developmental stages and in the different social interaction episodes targeted by the $PI^2$ dataset: face-to-face interaction, peekaboo, and reaching to objects in the presence of a caregiver.

**Modeling Caregiver Behavior** As explained in Section 1 the complexity of control problems can be measured in terms of the degree of presumed knowledge. In the simplest case (MDP), the controller is certain about the state of the world and the dynamics of the world. A higher degree of realism is achieved when the controller is given a model of world dynamics but needs to infer the current state of the world based on sensory information. In the third, most realistic case, the controller needs to learn a model of the world and the state of the world based on sensory information (BAPOMDP). In the case of early social interaction the world model refers to models of the caregiver’s behavior. For example such models may characterize the probability a caregiver will smile within 1 sec of a child smile. While the ultimate goal in this project is for the robot to learn models of caregiver behavior on its own, here we propose to use the $PI^2$ dataset to proceed by steps, developing first such models using standard statistical methods (e.g, the EM algorithm applied to partially observable Hidden Markov Models [16]). There are good reasons for doing so: (1) To compare the models learned by the robot on its own, with the models one could learn with classical statistical approaches using perfect sensors. (2) To allow experimentation with robots that are given different degrees of prior knowledge.

**7.4 Learning To Interact with People**

We propose to develop stochastic optimal control models for the CB2 robot to operate in three situations targeted by the $PI^2$ dataset: face-to-face interaction, peekaboo, and reaching to objects in the presence of a caregiver. The reward/cost function in these models will be based on the estimates developed using the inverse optimal control methods described above. An issue of interest will be to investigate the degree of prior knowledge necessary for the robot to operate in a realistic infant-like manner. First we will experiment with the case in which the robot is assumed to have already developed
a model of the world, including the operating characteristic of human caregivers. The models given to the robot will be those developed based on the $P1^2$ dataset, as described in the previous section. We will then progressively reduce the level of information in the prior models from fully known, to tabula-rasa. This can be elegantly done within the Bayesian framework by treating the models as transition probability matrices with Dirichlet priors. The parameters in these prior distributions can be varied to span all the cases of interest, from fully crystallized, to uninformative tabula-rasa.

Formally the problem we will be dealing with is of the BAPOMDP type, for which practical algorithms already exist [89]. The main limitation of these algorithms is that they require discrete states, sensor measurements, and actuators. The number of states can range between 500 and 1000, which is sufficient for our problem. The sensor information will consist of the CERT expression recognizer, touch detection, and vocalization detection. These will be discretized using standard vector quantization methods [36]. The complexity of the actuators will be progressively varied, starting with facial expressions and vocalizations and later adding arm movement and locomotion. We expect as part of this project for this team to develop new BAPOMPD algorithms capable of handling continuous states and observations.

Developmental experiments will be conducted in which human caregivers will interact with the robot on a daily basis simulating the 3 conditions of the $P1^2$ dataset. In all cases we will investigate the developmental sequence the robot goes through as it learns to interact with people. This developmental process will be compared to that observed in human infants. We will also measure success in terms of the realism of the observed human-robot interaction as evaluated by humans. We have extensive experience conducting this type of evaluations [79].

**Symbolic Communication: The Vigotsky reach.** Of particular interest will be the study of reach in the presence of caregivers. This problem bridges Section 6 where we study the problem of learning to control one’s body and physical objects, and Section 7 where we study the problem of learning to interact with others. In particular, early in development infants ignore caregivers when reaching towards desired objects. However by the end of the first year, typically developing (and not infants with autism) learn to use humans to obtain those objects. This process typically displays a a ritualization of a full reach gesture, which progressively becomes shortened into a reaching gesture. This gesturing, is typically accompanied by guttural vocalizations and looking back and forth between the caregiver and the object of interest. We call this pattern the “Vigotsky reach” in honor of the developmental psychologist that pointed out the important of this process [106]. The flagship target in this project is to have CB2 to learn and develop Vigotsky reaches on its own with as little pre-programing as possible. Developmental psychologists, like Vigotsky, have long focused on socially mediated reach as an example of the transition between sensory-motor forms of intelligence (the focus of this project) to symbolic forms intelligence (the focus of classical AI).
8 Management Plan

This project is a collaborative effort between machine learning and robotics faculty at the University of California, San Diego and the infant development lab at the university of Florida, Miami. The research team consists of three senior faculty members, two junior faculty members, six graduate students, and a postdoctoral student. The lead PI is Dr. Javier R. Movellan (Javier), a senior research professor at the University of California San Diego (UCSD). The PI from the university of Miami is Dr. Daniel Messinger (Daniel). The Co-PIs are Dr. Marian Bartlett (Marni), Dr. Emanuel Todorov (Emo), and Dr. Virginia de Sa (Ginny). Emo and Ginny are junior faculty members at the Cognitive Science Department at UCSD. Marni is an associate research scientist at the Institute for Neural Computation at UCSD. The project consultants are Dr. Hiroshi Ishiguro (Ishiguro), Dr. John Watson (John) and Dr. Terrence Sejnowski (Terry). Ishiguro is a very well known Japanese roboticist. His focus of interest is humanoid and android robotics. He was recently selected in the Synectics survey of contemporary geniuses, in company of people like Nelson Mandela, Stephen Hawking, Phillip Glass, Noam Chomski, George Lucas, and Steve Wozniak. John is an emeritus professor from UC Berkeley. He is a pioneer in the study of learning in infancy and its role in social development. In the 70’s and 80’s he generated many of the theories about early social development upon which this project is based. Terry is a senior scientist at the Salk Institute. He is one of the founders of modern computational neuroscience and a leader in the Machine Learning and Neuroscience communities.

In addition to the US team described above, this project would make possible a collaborative effort with a Japanese team, led by Dr. Minoru Asada and Dr. Hiroshi Ishiguro at the University of Osaka, Japan. The Japanese project is similar in goals to the one proposed here and is already funded by the Exploratory Research for Advance Technology initiative of the Japanese Science and Technology Agency. A critical tool in this project is a humanoid robot that approximates as much as current technology permits, the levels of complexity of human sensors and actuators. The Japanese team has already built a version of the robot, named CB2, and will help build a new replica of the robot to be hosted at UCSD. The robot will be built by the Japanese company Kokoro Dreams, LTD, the same company that built CB2, under the supervision of Ishiguro, Javier and Emo. Some small changes will be made to the current design, to improve its facial expressions and to make it approximate the body of a 1 year old infant rather than a child, as in the current version.

Javier will lead the project and will help coordinate the efforts of the UCSD team, the Miami team and the team at Osaka University. His training in behavioral and computational sciences and his life-long dedication to interdisciplinary research puts him in an ideal position to promote the interaction between the different components of the proposal. Dan will lead the behavioral aspects of the proposal, which will be based at the University of Florida, Miami.

An important aspect of this proposal is the collection of a dataset, named \( PI^2 \) of infant/caregiver interaction using motion capture and computer vision technology. This will be an unprecedented scientific resource that will help characterize the statistics of early physical and social interaction throughout the first year of life. To this end Javier, Emo and Marni, will assist John in the development of an Infant Motion Capture Facility at the University of Florida. They have already developed a similar facility at UCSD, albeit not specialized on infancy research. The facility will include an active motion capture system developed by Phase Space, and the CERT system for recognition of facial expressions from video developed by Marni. It will allow synchronous capture of caregiver/infant dyads, including head movements, trunk movements, arm and hand movements, and facial expressions (the last via the CERT system).

Dan will lead the collection of the \( PI^2 \) dataset and the additional behavioral studies proposed in this project. He will also conduct statistical modeling of the obtained data, as proposed in the main body of this document. Marni will focus on analysis and statistical modeling of infant/caregiver facial expressions with a focus towards understanding how they couple throughout development. She will also assist improving the morphology and the dynamics of the robot’s expressions so they approximate natural expressions produced by infants. Emo will lead the “learning to reach” effort as described.
earlier in this document. Ginny and Javier will lead the effort on “active discovery of object categories”. Javier will lead the effort on “learning to interact with people”, including the proposed studies to reverse engineer the cost/utility function infants optimize in early social interaction episodes, and the development of prior models of caregiver behavior using the proposed PI² dataset. Javier and Emo will work together towards the flag-ship problem in this proposal: developing a control algorithm for the robot to learn on its own to produce “Vigotsky reaches”. If this proposal is successful we expect the collaboration with the Japanese team to be very significant. Javier and Ishiguro have been collaborating since 2002, with Javier visiting his laboratory for extended periods of time. Three of his students (Joel Chenu, Nicholas Butko and Ian Fasel) have visited the Japanese group for several months. Some of the perceptual and learning primitives developed by Javier’s group have become part of his robots, and Javier has benefited from some of the robotics hardware developed by Ishiguro for his research. This project would help deepen this ongoing collaboration. Special emphasis will be made on student exchanges. We target for each year to have a US student to visit Japan for a period of 3 month and a Japanese student to visit the US for a similar period of time. We will also establish a joint repository of software and datasets focusing on developmental robotics. The repository will be made available to the research community and will include the proposed PI² dataset, as well as software developed as part of this project. In addition members of this research team will assist to the yearly symposium on synergistic intelligence which are currently organized by Dr. Minoru Asada. Two similar symposia will be organized by the US team, one to be hosted at UCSD and one at the University of Florida, Miami. Efforts will be made to invite K-12 Schools across the nation to interact with the researchers and the robot (physically and via Web) as the project progresses. To this effect a Web site will be developed by the Robotics Club at the Preuss School, with which Javier has been collaborating for the past 3 years. Preuss is a charter School for low-income student in grades 6-12. The demographics at Preuss 2004/05 were: 59.5% Hispanic, 21.7% Asian, 12.9% African American, 6% White. Last year more than 95 % of their graduating seniors moved on to 4 year Colleges around the country. Internship opportunities will be made available for Preuss students to participate in the proposed research activities.