Learning Complex Motor Tasks in Natural and Artificial Systems
A Proposal to the National Science Foundation

Investigators (all University of California, Berkeley, CA 94720):
Claire Farley (Integrative Biology)
Ronald Fearing (EECS)
Robert Full (Integrative Biology)
Richard Ivry (Psychology)
Stuart Russell (EECS)
Shankar Sastry (EECS)
Michael Dickinson (Integrative Biology)

A Project Summary

This project will develop a unified theory of how natural and artificial systems can learn to solve complex motor tasks, such as running, throwing [[RI]], and flying, that entail significant sensory input and the coordination, sequencing, and fine-tuning of many low-level activities. Such a project is possible because of significant experimental advances in our understanding of motor control systems in humans and other animals, and because of increased sophistication in our mathematical models of control learning. These models will be used not only to analyze and predict natural phenomena in motor control, but also to derive effective adaptive controllers for artificial systems carrying out complex tasks. In particular, a microelectromechanical winged insect will be trained to fly.

To generate complex behaviors, natural and artificial systems must be organized hierarchically with multiple layers of abstraction. The first research task will therefore be to identify appropriate levels of representation at which the physical system can be modelled and at which control actions can be defined. For example, in describing an insect flying from A to B, possible levels might be 1) nerve signals and mechanical properties controlling the detailed shaping of each wingbeat 2) basic wingbeat cycle 3) "steering" the cycle to direct flight 4) takeoff, navigation, landing. Detailed motion, force, and airflow measurements will be made under a variety of experimental circumstances and flying tasks to establish the correspondence. Similar experiments will be carried out for running in cockroaches and for running and throwing [[RI]] in humans. These studies (and, in the case of insects, neurophysiological studies) will also establish the sensory inputs that are available at each level of the control system.

Given the general structure of the control system and the appropriate sensory inputs, the next step is to design learning algorithms capable of learning to perform the given task successfully. The learning method to be used is *reinforcement learning*, a technique designed to adjust the control algorithm to optimize a specified *objective function*. The objective function is supplied to the learning algorithm as part of the sensory input in the form of a *multiattribute reward signal* at each point in time—the objective function being some additive function of the rewards over time. New reinforcement learning algoriths will be developed that operate using both local and global reward signals within a hierarchical control structure; furthermore, these algorithms will be proved to converge even using nonlinear representations of the objective function.

It is hypothesized that natural control systems may be implicitly optimizing an objective function, and that some systems may use a form of reinforcement learning from a reward signal to do so. It is not known what the reward signal might be, but the failure of previous simpler proposals suggests that it must be a complex, nonlinear, multiattribute combination of the sensory inputs. The final step in establishing correspondence between natural and artificial systems is therefore to ascertain what it is that the natural control system is implicitly optimizing, if anything—that is, to "reverse engineer" the reward function by fitting the mathematical model to observed behavior in a variety of circumstances. This task is a new form of computational learning that has not been studied before, although it relates to the ascertainment of multiattribute utility functions in economics. The team will derive algorithms for this task and will show mathematically the circumstances under which such an ascertainment can be made. They will then carry out an experimental program for the tasks and systems described above, including the use of ascertained reward functions to predict system response under new circumstances. This research should shed light on the central question of whether this form of learning in animals and humans can be viewed as driven by optimization or by some other principle, such as the preservation of fixed interface characteristics among the various levels of the system. Discovery of consistent reward functions in animals, especially humans, would have significant consequences for general theories of learning.

C Project Description

C.1 Objectives, expected significance, longer-term goals

[[rewrite to reflect summary; stress common language—for bones and muscles, mechanics (mass, length, force etc) is a common language, but for AI/EE/Bio control systems there are (so far) no agreed-upon abstractions.]]

The principal scientific aim of the proposed research is to develop a unified theory of how intelligent systems (human or artificial) can learn to solve complex motor tasks, where the target performance entails significant sensory input and the coordination, sequencing, and fine-tuning of many low-level activities. Scaling up the learning task in this way acts as a forcing function that distinguishes between viable and nonviable approaches to learning, in a way that simple tasks cannot. Understanding and solving such complex learning problems will involve a concerted effort from the skill sets of all our team members—artificial intelligence, control theory, reinforcement learning, psychology, biomechanics, and neurophysiology.

All learning methods are essentially techniques for searching in a space of system configurations in order to find a 'good' configuration. Without imposing some additional structure, any learning method will, in the worst case, take time and experience that are exponential in the "size" of the configurations being considered. (This is true for both evolutionary and online learning.) To date, the only universally applicable solution, one that has emerged in several disciplines independently, is the use of some form of hierarchical or multiscale decomposition. Our hypothesis is therefore that complex motor learning tasks inevitably involve hierarchical and/or multiscale sensing, modeling, and control; in turn, this has specific implications for the range of possible learning methods.

As yet, sensorimotor hierarchies are not well-understood theoretically. We can (in some cases) design and verify specific hierarchies, but there are many fundamental issues remaining open, such as: What is the right decomposition? How can it be learned? How can components within a given decomposition be optimized based on a global reward signal? Where should the locus of control be during "execution" of hierarchical control strategies? How robust are learning methods for hierarchical models and policies? How can specific hierarchical strategies be generalized and adapted to new situations? How can specific skills be learned, abstracted, and recombined into new complex behaviors? What communication needs to occur in distributed sensing and control systems? How can the control strategy accommodate online revisions in the face of new sensor data?

To answer these questions, we will develop specific mathematical models and learning algorithms for hierarchically structured, hybrid (discrete+continuous) control systems. We will derive analyses of convergence, learning rates, complexity, and robustness.

[[more here from summary: unification research agenda: The next step is to provide the basic formal relationship between the empirical data and mathematical models of learning; this requires identifying the sensory input data available for learning (using neurophysiological studies of flies and anatomical knowledge of humans), and identifying the *objective function* that defines the direction of learning. Mathematical models will be devised for hierarchical control based on reinforcement learning, and algorithms will be constructed that provably optimize nonlinear objective functions within these models. Given such algorithms, it should be possible to "reverse engineer" the objective function that an animal may be optimizing as it adapts its control system to changing circumstances. The team will show mathematically the circumstances under which such an ascertainment can be made and will propose and carry out an experimental methodology based on predictive modelling. This research should shed light on the central question of whether this form of learning in animals and humans can be viewed as driven by optimization or by some other principle, such as the preservation of fixed interface characteristics among the various levels of the system.]]

A strong synergy is expected between this analytical work and psychological and physiological experimentation. At present, human abilities to acquire complex skills from very limited training signals far exceed those that can be explained by our neurophysiological models or those that we can implement in artificial systems. The same is true even for some "hard-wired" control systems, such as that employed by the hoverfly. Our psychological work on "gating networks" and neurophysiological evidence of the decentralization of motor control may therefore suggest new (possibly even nonhierarchical) approaches to the design of artificial learning systems. Conversely, we have several new models, arising from our work in reinforcement learning and hybrid control systems, that suggest new experimental investigations at the psychological and neurophysiological level. More importantly, they will provide new ways of describing and understanding the empirical findings in such a way that they can be mapped onto analytically tractable models of control systems.

The members of our team have all participated in the development of new approaches to the modelling of complex behavior and learning in their respective disciplines. In the following sections, we explain these contributions and sketch the beginnings of a unified approach. We also describe two testbed projects, for which we have already developed significant expertise and infrastructure:

1. Running in cockroaches.
2. Running and throwing [[RI]] in humans.
3. Flying in both artificial insects (silicon/piezoelectric microelectromechanical systems) and real insects (hoverflies and fruit flies).

The proposed research will have long-term impact in several ways:

- *Scaling up intelligent systems*:
- *New control mechanisms for complex physical systems*:
- *Detailed models for complex motor learning in animals and humans*:
- *Better teaching, training, and rehabilitation methods for physical skills*:

A more abstract but equally important contribution will be a unified *vocabulary* for describing the organization and adaptation methods of complex control systems—a vocabulary that is interpretable by scientists in biology, control theory, and artificial intelligence and enables useful collaboration.

C.2 Learning complex behaviors

When we (or nature) are designing a system for learning complex behaviors, certain key questions must be answered:

1. What is the functional organization of the control system that will carry out the complex behavior, and what parts are amenable to modification through learning?
2. What feedback is available to indicate the directions for improvement?
3. What adaptation mechanisms can be implemented to effect the necessary modifications based on the feedback?

A fourth question arising in the design of artificial systems is that of how to combat complexity, both in the computations required for control and the amount of experience required to learn behaviors. Nature has already solved this problem, and we derive some inspiration from her designs.

In each of the fields represented by the members of our project team, various ideas have been proposed to answer these questions. In this section, we discuss those ideas, providing the necessary background to make them comprehensible. We lay out some basic research questions and make proposals for answering them.

[[EACH OF THE FOLLOWING SUBSUBSECTIONS DESCRIBES WHAT WE'VE DONE, WHY OUR WORK IS TRIFFIC, INCLUDE RELATED WORK BY OTHERS. DOESN'T NECESSARILY NEED TO ANSWER THE ABOVE QUESTIONS 1,2,3 BUT THEY MAY HELP IN MAKING THE SECTIONS COHERENT]]

C.2.1 Adaptive intelligent control

The title of this subsection reflects its origin: a synthesis of ideas from artificial intelligence (particularly reinforcement learning) with ideas from adaptive optimal control theory. This synthesis is part of an ongoing project at Berkeley to find common ground between these two fields. We begin by providing some background and establishing some interdisciplinary terminology.

Controllers

A *controller*[1] is simply a device (computational, mechanical, or neural) that receives sensory inputs and generates outputs that affect the physical world (in control-theoretic terminology, the *plant)*. Depending on the task definition, controllers can range from extraordinarily simple (a thermostat) to extraordinarily complex (a brain and nervous system). In this proposal, we will be systems of moderate complexity for physical tasks, since these offer the best hope of finding a real and fruitful correspondence between theoretical constructs and actual biological systems, with consequent benefits for the design of artificial systems.

We will use the term *policy* to denote a complete input/output description of a controller, i.e., a policy defines the controller's response to all possible input sequences.[2] For any given policy, many controller designs are possible. One useful classification of controllers is into *model-free* and *model-based*, where the term "model" refers to some representation of the *effects* of the actions and the dynamics of the environment (see below).[3] Model-free controllers can be *direct* implementations of a policy—for example, a neural network whose inputs are sensor measurements and whose outputs connect to the physical effectors.[4] A second kind of model-free controller represents the policy using a *Q-function* (Watkins, 1989), which gives the *value* of taking each action in any given state; then some kind of comparison mechanism must select the action of highest value for the current state. (See also Section C.2.2.)

Notice that neither of the model-free designs A model-based controller, on the other hand, can use the model to predict the outcome of each action. Given some representation V of the values of states, it can evaluate those outcomes and decide what to do. This is the basis of a wide range of controllers, ranging from *model-predictive control*, widely used in industrial processes (Nevistic & Morari, 1996), to chess-playing programs in AI.

Models and model learning

In control theory, a model is typically a set of differential equations, stating how the state will evolve over time depending on the control action applied—for example, what are the angular and longitudinal accelerations of a car as a function of the steering angle and throttle value (and other state variables)? AI, on the other hand, typically deals with discrete states and time steps—e.g., chess moves can be described using logical assertions.

A model can be *deterministic* or *stochastic*, depending on whether there is uncertainty about control outcomes. (For all real systems, any deterministic model is an approximation, so much of control theory is concerned with errors and perturbations.) Stochastic models in control theory include stochastic differential equations and linear models with Gaussian noise. In AI, dynamic Bayesian networks provide a very general stochastic modelling language.

Learning a model means identifying equations/rules/networks that correctly predict the next state from the current state and action. When the environment state is *observable*, model learning can be fairly easy if we know the *form* of the model, e.g., for the human body as a set of differential equations with unknown parameters such as masses, limb lengths, moments of inertia, etc. These parameters can typically be learned from observing the body's response to control signals.

Regulatory and optimal control

[[SS–couple of sentences abut regulatory control? It's a significant alternative hypothesis to optimal/RL is it not?]]

Optimal control, as its name suggests, involves maximizing some *performance measure*.[5] For example, in driving, we might want to optimize some combination of speed, safety, legality, passenger comfort, fuel efficiency, etc. An optimal controller is one that maximizes the expected performance measure compared to all controllers in some defined class.

Direct optimization of control parameters

The classical theory of optimal control (Athans & Falb, 1966) assumes a deterministic model in the form of differential equations, and, given certain assumptions on the nature of the model and the performance measure, can derive equations defining an optimal controller. The theory now encompasses adaptive control of certain nonlinear systems (Sastry & Bodson, 1989) and has a long string of impressive achievements in practice.

[1] In AI terms, an *agent*.

[2] In AI, this full is called the *agent function* (Russell & Norvig, 1995).

[3] [[SS-Is this OK?]] One important difference between AI and control theory concerns *who has the model*—the designer or the controller itself. Here, we are concerned with the controller.

[4] One well-known instance of this form of controller is the ALVINN system for vision-based driving (Pomerleau, 1993).

[5] We deliberately choose this neutral term to avoid confusion. AI uses *goal* and *utility function* and control theory uses *objective function*. Biology uses [[what? fitness function? CF]]

Unfortunately these achievements do not come close to the difficulty of motor tasks that humans learn every day. We routinely learn to perform complex motor actions such as walking, running, riding bicycles, gymnastics, diving, writing, speaking, and windsurfing. If one were to write mathematical models of these activities, they would involve complex dynamical equations of multi-body systems (the limbs of the body) interacting with complex contact constraints (running on sand, uneven terrain) and with other multi-body systems (the bicycle), subject to non-holonomic conservation laws (the diver or gymnast) and complex aerodynamic, hydrodynamic forces (the windsurfer). For example in our work on diving (Crawford & Sastry, 1995, 1996), we have found that even a grossly simplified 10-degree-of-freedom model of the body (3 degrees of freedom in each shoulder, 1 in each hip, and 1 in each elbow) produces equations of motion covering several pages in symbolic form—and this is the simplest case of motion because there are no contact points or forces. Thus, in general, we cannot hope to solve optimal control problems in closed form even if an exact model can be learned.

[[SS–say something about various ways around this OTHER THAN reinforcement learning, eg learning the weights of a direct neural net controller, dating back to Widrow?]]

[[SS–not sure how to categorize these: detailed biological approaches such as the Adjustable Pattern Generators modeled on the cerebellum (Berthier, Singh, Barto, & Houk, 1993). Many fall somewhere in between these, often with neural networks as part of the dynamic control structure (see (Agarwal, 1997) for a review), such as Kawato's feedback error learning scheme (Kawato, 1990). Recursive learning based on continuous, parametrized controls has been used with radial basis functions to train a space robot with a nonholonomic arm (Gorinevsky, Kapitanovsky, & Goldenberg, 1996), as well as with a Jacobian estimate to optimize the design of a robot arm (Pil & Asada, 1994).]]

Reinforcement learning

In recent years, *reinforcement learning* (also called *neurodynamic programming*) has made rapid progress as an approach for solving complex decision and control problems (Sutton, 1988; Kaelbling, Littman, & Moore, 1996; Bertsekas & Tsitsiklis, 1996). The basic idea is that the performance measure is made available to the adaptive controller in the form of a *reward function* specifying the reward for each state that the controller passes through. The performance measure is then the sum of the rewards obtained. For example, when a bumble bee forages, the reward function at each time step might be some combination of the distance flown (weighted negatively) and the nectar ingested.

A model and a reward function together define a *Markov decision process* (MDP), whose solution is a policy that maximizes the expected sum of rewards. An optimal policy induces an optimal *value function* $V^*(s)$ on states s, giving the total expected reward from that point on if the optimal policy is followed. For the discrete stochastic case (studied in AI and operations research), let the model be specified by the probability $P(s'|a,s)$ of reaching state s' by taking action a in state s. Then the optimal value function satisfies the *Bellman equation* for each state (Bellman, 1957):

$$V^*(s) = R(s) + \max_a \sum_{s'} P(s'|a,s)V^*(s')$$

where $R(s)$ is the reward for s. In the continuous deterministic case (studied in control theory), we have the differential Hamilton-Jacobi equation: [[SS–all yours]]

These equations can be solved offline to give an optimal value function (and hence an optimal model-based controller); alternatively, analogous equations can be solved to find $Q^*(a,s)$, the long-run value of doing a in s, which can be used in a model-free controller. However, the solution process is computationally infeasible for most real systems whose state space is typically vast or infinite, and cannot be attempted if the model is unknown or error-prone.

Reinforcement learning avoids these problems by learning directly from the observed rewards while acting in the environment. An estimate $\hat{V}$ of V^* can be learned using the *temporal difference* (TD) update rule (Sutton, 1988), and an estimate $\hat{Q}$ of Q^* can be learned using the Q-learning rule (Watkins, 1989). The complexity is proportional to the number of states visited. The learned estimates can be extended to all states by representing the estimated functions using a *function approximator* such as a neural network. In this case, the update rules modify the network parameters directly. This approach has proved successful in practice; for example, the world's best backgammon player is a program trained from scratch using TD (Tesauro, 1992). We hypothesize that 1) reinforcement learning can enable an artificial system to learn the kinds of complex motor behaviors we are studying; 2) some analogue of reinforcement learning plays a role in adaptive motor control in animals.

Fundamental problems in reinforcement learning

Convergence: For the discrete case, with anything more powerful than a linear function approximator (Tsitsiklis & Van Roy, 1996), TD learning is not known to converge. In fact, divergence and "unlearning" are serious problems in practical applications of TD with neural nets. We have developed a new family of algorithms based on taking linear combinations of successive estimated value functions, and have preliminary results (Papavassiliou & Russell, 1998) showing that

- convergence is guaranteed for general function approximators;
- convergence to a *global* optimum is guaranteed if the function approximator family is agnostically learnable (as defined by Kearns, Schapire, , & Sellie, 1992);
- convergence is efficient if the underlying Markov chain mixes rapidly.

We plan to further develop the analysis of this new algorithm, and to apply it to control problems in the near future.

For motor control, we must consider convergence of reinforcement learning in continuous domains, for which almost nothing is known. With recent advances in fast marching algorithms for Hamilton Jacobi equation calculations (see for example (Tomlin, Lygeros, & Sastry, 1998)), we feel that we will be able to give more definite proofs of convergence of methods for approximate solutions in the continuous deterministic case. While this body of literature, originating in aircraft control, is seemingly distant from the motor learning proposed in this project, conceptually it is clear that the two fields share similar analytical methods for proving convergence results.

Partial observability: [[SR: explain the problem; V(belief state); initial results (Parr & Russell, 1995). Approach based on DBNs (Friedman, Murphy, & Russell, 1998); baseline architecture]]

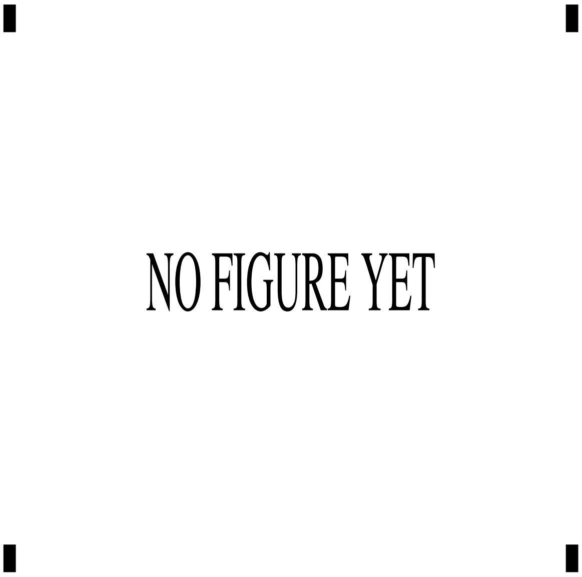
Scaling up: Even with function approximation etc., simple reinforcement learning of a single value or *Q*-function from scratch is jsut not powerful enough to handle difficult control problems, especially with many degrees of freedom and/or complex sequences of actions required. Hence, in the follwong paragraphs, we propose a solution approach based on a *multi-level hierarchical* approach. In Section C.3, we discuss further research needed to complete the picture.

A multi-level adaptive controller [[SS: discuss Bernstein (1967) (old hat but relevant), inspiration from biological control systems (other refs?), basic ideas including virtual plant (better name?)]]

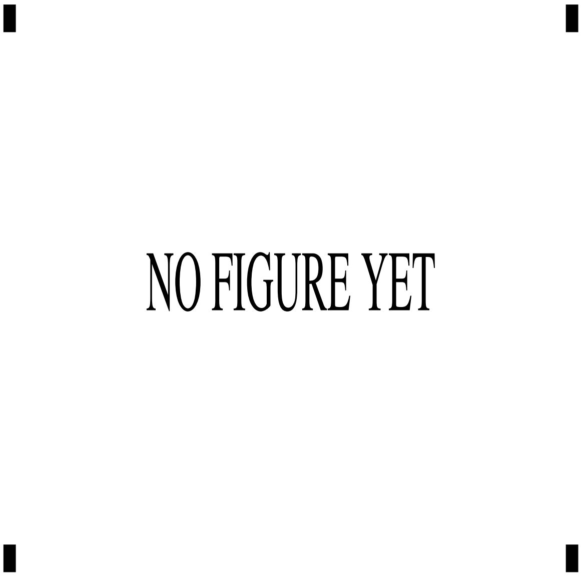
The controller design presented here, shown in Figure 1(a) for the example of diving, was inspired by the hierarchical, behavioral control structure used by biological systems use to control their movement. The dives executable by the system are represented as a parametrized class; the parameters specifying a particular desired dive then form the input to the controller. The high-level coordinating controller coordinates between several multi-joint behavioral synergies, learning the timing between them. The multi-joint synergies themselves provide the tuning inputs for the low-level single-degree-of-freedom (single-DOF) controllers, each of which plays the role of a pattern generator for fast, single-joint movements. The coordinating controller and the single-DOF controllers all operate in the discrete time domain. The plant, a mechanical system, operates in continuous time; the pattern generator structure in the single-DOF controllers allow this simplification of the control representation and provides the link between the two regimes.

A closeup of a low-level single-DOF controller is shown in Figure 1(b). This is inspired by the design of Gottlieb, Corcos, and Agarwal (1989) for a controller for fast, goal-directed, single-joint movements, which can be seen as a pattern generator for these kinds of movements. Because of the many degrees of freedom and because the desired dive is only specified by general, behavioral metrics, there may be many possible control torque profiles that would achieve the goal. A pattern generator is one way of restricting these control choices to torque profiles within a certain parametrized family, making the control problem easier. As shown in Figure 1(b), the single-DOF controller is made up of a control network, a decoder, an inertial compensator, and a filter.

The control network takes as input a vector $[y_d, \dot{\theta}_0] = [\Delta\theta_d, \Delta\dot{\theta}_d, T_d, \dot{\theta}_0] \in \mathbb{R}^4$ describing the desired change in the DOF angle (a particular Euler angle, if the DOF is part of a multi-DOF joint), the desired change in joint velocity, the desired movement duration, and the current DOF velocity (sensory information available from the spindle organs in biological systems). The output of the network is a vector in $\mathbb{R}^3$ giving a parametrization of the control torque profile.



(a)



(b)

Figure 1: (a) Overview of the hierarchical design for the learning controller applied to the learning of a dive. (b) Design of the single-DOF controller.

The torque profile family used is a filtered version of a two-rectangular-pulse function, as shown in Figure 1(b). The control network thus provides the translation from kinematics to dynamics, all in a discrete time regime, abstracting the dynamics out of the control and learning problem; the controller network need not provide continuous-time torque information directly. By restricting the torque to a parametrized family of functions, the single-DOF controller defines a behavioral model for the allowed class of single-DOF motions.

A closeup of the coordinating controller architecture for the instance of diving is shown in Figure 2(a). Based on information the diving student would obtain from an instructor, the dive is broken up into several parametrized behavioral synergies. From instructor input and from watching other divers, the student learns the order in which these behaviors must be performed and an initial estimate of the timing required between them. The controller therefore behaves like a timed state machine as it switches between behavioral synergies, with a different state machine for each learned dive. The controller must then refine the parameter estimates learned in the "learning by watching" stage; for this "learning by doing" stage, the Q-learning algorithm may be used with a function approximator. Other RL schemes to be developed on this grant (see above) will be used as we develop greater convergence knowledge of these methods. The coordinator network's model is updated at each transition as well, using its own estimate of the goodness of the new state as feedback for learning. A flow chart showing the linked processes of action selection, simulation, and learning, is shown in Figure 2(b).

Additional flexibility: Learning complex behavioral sequences and skills

The controller for diving described above assumes a single, fixed behavioral sequence and learns when to switch each behavior in and out and how to execute each behavior to maximize the total reward. This is *parametric* flexibility. Humans, however, can learn new behaviors whose sequential composition varies with the situation—for example, we can clamber over rocky terrain or drive in heavy weaving traffic. Here, we propose a flexible structure for behaviors and give some basic results.

[[Define Hierarchical Abstract Machines (HAMs)
state theorems for basic version of offline and online learning
discuss generalization
learning HAM structure]]

[[From 1997 LIS proposal:
By combining work in control theory, model learning, and AI reinforcement learning, we are developing a mathematical model for the learning of complex control strategies that captures all of these features. The key aspect is learning *hierarchical* strategies that are *flexible* and *context-sensitive*. Recent work by Parr and Russell (*ML-97*) has shown

NO FIGURE YET

(a)

NO FIGURE YET

(b)

Figure 2: (a) Design of the coordinating controller. (b) Flow chart showing the relationships between the processes of action selection, simulation, and learning, for the instance of diving.

that this can be achieved within the formal framework of reinforcement learning and Markov decision processes, via the specification of hierarchies of controllers with undetermined choice points (such as which motions to execute in a dive) that define an abstracted, exponentially smaller, MDP. Initial tests have shown dramatic increases in the speed of learning, with "leaps" in performance as new levels of the hierarchy are learned. This research promises to unify, explain, and extend several lines of work by Brooks and others on layered models of behavior, and to yield learning systems of greater power than current schemes.]]

Control with concurrent subsystems
Many high-level skills can be seen as being comprised of a few major sub-behaviors, running in either series or parallel. For skills such as diving that involve a series of linked sub-behaviors ("throw," then "pike," then "layout," for example), the Markov decision process framework required for standard reinforcement learning algorithms is a reasonable one. For other skills that involve simultaneous, parallel sub-behaviors, however, a Markov decision process model is extremely poor, since the "outcome" states occur asymchronously. In walking or running, for example, three subroutines run in parallel: the support, balance, and swing control synergies (Winter, 1995). For the support moment synergy, the muscles in the legs work together during the stance phase of walking to keep the body from collapsing. The kinematics of walking are remarkably similar from one stride to the next but the motor strategy used to achieve these kinematics can vary substantially from stride to stride. Thus, global kinematic patterns are conserved while the activation of individual muscles and the movement of individual joints varies. Changes in one muscle's action or one joint's movement are offset by other muscles or joints. For the balance synergy, the muscles in the hips work to keep the torso upright. The swing control synergy is responsible for the swing and placement of the swing foot. These parallel, interacting, synergetic behaviors produce the overall behavior of walking or running. We propose to research extensions of the hierarchical reinforcement learning schemes described above to handle this case. We will use the theory of *semi-Markov* decision processes, which allow actions with continuously varying duration (see, e.g., ?). We believe this theory can be augmented further to handle parallel asynchronous composition of actions.

[[SR: ideas for how to scale up AI/control ideas to hard problems; input generalization for HAMs; local reward theory (summation constraint)]]

C.2.2 Basic psychological findings

Motor control and learning have tended to be neglected topics in cognitive psychology. Part of this stems from the fact that motor control has relatively little to do with the mind and more to do with the muscles. Thus, our intuitions suggest there are fundamental differences between what underlies individual differences in motor skills compared to non-motor areas of skill. There is surely some truth in this idea. People who excel in sports are frequently not the best achievers in non-athletic domains and the reverse if also obviously true. Paralleling the concept "intelligence", the most basic term we use in the motor area is "coordination" to characterize gross differences between individuals. When we see someone performing at a world class level in baseball, basketball, or tennis, we talk about how coordinated the person is, how graceful and effortless their performance appears to be. The concept of coordination immediately conjures up images of a well-tuned skeletomuscular system. A corollary of this idea is that much of what separates them from the rest of us Saturday afternoon jocks is to be found by examining their neuromuscular or biomechanical capacities. Their muscles are bigger, respond faster or, at the very least, in a more consistent manner. They have more efficient metabolic systems for transforming the raw ingredients contained in air, food, and water into the energy sources required to drive performance.

What is frequently missing in these intuitions is the brain. And yet when we look at the literature on motor skills, what is most striking is the similarity between how we account for skilled performance in athletics and how we account for elite performance in more cognitive domains (e.g., Ericsson, Krampe, & Tesch-Roemer, 1993). It appears that much of the difference between elite and non-elite performers in motor skills has little to do with the muscular system. Some of the evidence here comes from studies that are modeled on those used to examine cognitive skills. In their classic studies, Chase and Simon (1973) investigated the differences among chess players, asking what is the best way to characterize the advanced skill of players who had reached the master level. Using a variety of memory recognition tasks, they showed that masters appeared to have an enormous memory store of possible chess positions. Most important, they found that this enhanced knowledge was specific to chess; it did not reflect an inherent difference in some sort of general memory capability, either in terms of the amount of information that can be encoded or the rapidity with which that information is retrieved. The chess masters were not better at memory tasks in general nor in tasks using chess pieces that were randomly arranged on chess boards (and thus not part of their long-term knowledge of chess). Similar studies have been carried out in a variety of motor skills and the results here also indicate that elite performers have tremendous pattern recognition capabilities that are specific to their domain of expertise (Allard, Graham, & Paarsalu, 1980). Skilled basketball players gain an important advantage over their opponents by being able to recognize the position of their teammates and opponents. Magic Johnson's great passing skills need not be attributed to his quick hands or reflexes, but rather to the fact that his memory advantage allows him to efficiently retrieve the motor plan that is most appropriate for a particular context.

It is unlikely, though, that we are going to account for individual differences in motor skills purely by focusing on the advanced perceptual abilities of elite performers. Nonetheless, when we study the development of motor skills, we find that much of what is learned can not be ascribed to the level of muscles or peripheral factors. The concept of the "motor program" has been proposed to account for the knowledge structure guiding the performance of skilled action. A motor program is a representation of the movement commands that underlie a complex gesture. A central concern in this area of research has been to identify the level at which these commands are specified. Actions entail hierarchic representations in which the movements are initially specified in a relatively conceptual manner (i.e., the goal to be achieved) and then translated into a particular pattern of muscular activity. For example, the hitter in baseball is intent of achieving a certain configuration of the arms and hands when contacting the ball. The exact movement pattern required to achieve this configuration will vary depending on the initial posture as well as the speed and location of the ball. Learning occurs at all levels in the hierarchy (MacKay, 1987; Wright, 1990). Surprisingly, studies of motor transfer tend to show that much of what is learned is at the more abstract levels. Thus, when someone is trained to produce a series of sequential movements with the fingers, they show near perfect transfer when the task is changed so that the movements are now made with whole arm movements (Cohen, Morgan, & Pollack, 1990). Learning must be relatively abstract since different muscles are required to move the hand in comparison to the fingers. One caveat here: Most of these studies have involved fairly simple movements in which the lower levels of the hierarchy (e.g., the activation patterns across the muscles) are well established prior to training. A different pattern of results might emerge if similar studies were carried out requiring more complex movements.

Another aspect of skilled performance is that learning entails a transition from early stages in which extensive cognitive effort is required to advanced stages in which performance becomes more "automatic." The term "automatic," however, has many different senses. One notion is that the degree of control is reduced. However, studies of expert performers tend to show the opposite in that they are more sensitive to errors or unexpected perturbations than novices (Logan, 1982). An alternative conceptualization of automaticity is that over the course of skill acquisition, the viability of different response alternatives is modified. Early in training, a context may lead to the activation of many possible actions and relatively extended processing time is required for one of these actions to be selected (Ivry, 1997). With practice and reinforcement, the system becomes biased to favor a particular response or small set of alternatives within a particular context. This idea might be instantiated as a change in the matrix of Q values, with a shift from little to large differentiation between the set of values as skill develops.

[[FOLLOWING IS MODIFIED FROM CURRENT DRAFT. BUT IT MAY NOT FIT IN ANYMORE WITH ABOVE SECTION AND GENERAL EMPHASIS ON LEARNING IN THE PROPOSAL]]

Planning and selecting a particular action are two central operations performed by a cognitive system. There also remains the problem of how a particular action is implemented. Our experimental research on bimanual movements (Helmuth & Ivry, 1996; Franz, Eliassen, Ivry, & Gazzaniga, 1996) and multi-task performance (Ivry, Franz, Kingstone, & Johnston, 1998) supports the idea that a fundamental constraint in motor control arises from a winner-take-all gating cycle, in which candidate actions compete for execution, with the winner emerging via a competitive network. The consequences of learning will be at least two-fold here. First, learning will influence the content of the action "chunks" that are gated. Presumably, early in learning, the chunks are small and a task requires many gating cycles. With experience, the amount of information per cycle becomes larger. Thus, the novice pianist will hunt and peck for each note whereas a series of notes are produced as a unit by the expert. Second, the number of candidate actions achieving significant activation in the gating cycle as part of the "automatization" process described above. Both of these hypotheses are consistent with the algorithmic picture sketched above. A final hypothesis is that the gating mechanism acts as a central controller to enusre the coordination of actions and minimize the potential for conflict (Berns & Sejnowski, 1996). An implication of this hypothesis is that there are limits on how well people can perform two actions simulataneously, even if these actions involve different input and output systems (Pashler, 1998). For example, one cannot drive and talk at the same time without interference, although the detection of such interference may require subtle measurement.

C.2.3 Basic biological findings

Natural systems show great adaptability and robustness in motor control tasks such as walking, running, and flying. Much of the robustness is inherent in the structure and morphology of the organism, as shown by results from humans, cockroaches, and flying insects. Low level sensorimotor control loops seem to strongly depend on abstract locomotion principles inherent to the structure and environmental interaction. We can consider evolutionary adaptation of structure and sensorimotor control loops as a form of programming the organism how to interact with the environment.

The dynamical dependence between seemingly unconstrained degrees of freedom makes understanding the control and learning of complex motor tasks a true challenge. Although neuro-biological systems hold great promise for building adaptable artificial systems, classical circuit breaking techniques in neurobiology of even the simplest systems have seldom resulted in the neural models where tuning, switching or learning strategies are obvious and can be linked to behavior. The reason is becoming increasingly clear. (Raibert & Hodgins, 1993) stated it this way, "Many researchers in neural motor control think of the nervous system as a source of commands that are issued to the body as direct orders. We believe that the mechanical system has a mind of its own, governed by the physical structure and laws of physics. Rather than issuing direct commands, the nervous system can only make suggestions which are reconciled with the physics of the system and task at hand." Mochon and McMahon. (1980, 1981) showed the importance of a completely passive leg swing in ballistic walking. Playter and Raibert (1994) demonstrated how movable arms projecting out from a body can stabilize rotation by only using springs. McGeer (1990) and colleagues have built downhill walking machines which have no sensors or actuators and lack any computer control. For years, Raibert devised spectacular hopping, running and trotting robots with only very simple control systems (Raibert & Hodgins, 1993). Full (1993) has shown that despite spectacular diversity in animal locomotor morphology, rapid legged locomotion can be characterized by a single mechanical model - a spring-mass system similar to the Raibert hopping robot (see Full & Farley, 1998). From centipedes to cockroaches to dogs and even humans, animals that differ in leg number, posture, and body form appear to have one value for relative leg stiffness (Blickhan & Full, 1993). Before this discovery the control problem appeared complex, varied, and devoid of simple higher level structures or "suggestions."

Nature tells us that mechanical imperatives of dynamic systems can greatly simplify the structure of the control system (Ting, Blickhan, & Full, 1994). We must realize that musculoskeletal and mechanical systems can also do computations - passive, yet dynamic computations. In these passive dynamic systems, the algorithms are simply embedded in the form of the machine or animal itself. Control results from the properties of the parts and their morphological arrangement. In animals, musculoskeletal units, leg segments and legs do much of the computation on their own by using segment mass, length, inertia, elasticity, and dampening as "primitives". These passive dynamic controllers can respond immediately (i.e., zero order) and effectively to a variety of perturbations.

Human Running

The relative roles of passive and active control are starting to become clear in studies of locomotion. Running can be modeled very simply, using a spring-mass model (Alexander, 1988, 1992; Blickhan, 1989; Blickhan & Full, 1993; Farley & Gonzalez, 1996; Farley, McMahon, & Glasheen, 1993; Ferris, Louie, & Farley, 1998; He, McMahon, & Kram, 1991; McMahon, 1990; McMahon & Cheng, 1990; Thompson & Raibert, 1989). This spring-mass model has been shown to describe and predict the mechanics of running gaits remarkably well. According to this model, the dynamics of the center of mass during the ground contact phase are completely determined by the stiffness of the leg spring ("leg stiffness"), the touchdown angle of the leg, and the touchdown velocity.

Evidence is building that humans and other animals rely heavily on passive dynamic control during running. Under steady-state conditions, running humans adjust leg stiffness to accommodate different surface stiffnesses. The extent of leg stiffness adjustment is such that the total stiffness of the series combination of the leg and the surface is always the same, regardless of surface stiffness (Ferris & Farley, 1997; Ferris et al., 1998). As a result, many aspects of running remain the same on different surfaces, including the peak ground reaction force, ground contact time, and the total vertical displacement of the center of mass.

Recently, we have begun to examine transitions from one surface to another. If a runner is not aware that the change in surface stiffness is going to occur, then leg stiffness is not adjusted in time for the first step on the new surface. In this situation, the movements of the body's center of mass are exactly as predicted by the simple spring-mass model. As an example, consider the case of a person who is running on a compliant surface and is surprised by a sudden transition to a very stiff surface. In this case, the person's leg is too stiff for first step on the extremely stiff surface. As a result, the person slows down substantially after the foot hits the ground, and their leg angle at take-off is not the same as their leg angle was at touchdown. As a result, the person appears to be unstable in the fore-aft direction for the next few steps until they eventually find the right combination of leg stiffness and touchdown leg angle for that particular surface. There is no evidence of any effective neural intervention during the first ground contact phase (280 ms) after a surprise transition to a new surface. These observations strongly suggest that the runner's movements during a given ground contact phase are mainly determined by their preset leg stiffness and touchdown leg angle as well as the initial velocity of the body's center of mass. Within a given ground contact phase, this form of "passive dynamic control" dominates with very little active neural control.

If a person is given the opportunity to practice making the transition from one surface to another, they can rapidly learn to preset their leg stiffness so that it is perfectly adjusted in time for the first ground contact phase on the new surface (Ferris et al., 1998). Thus, via learning, they become able to adjust leg stiffness in the time between their last footfall on one surface and their first footfall on the new surface. After this learning has occurred, the movements of the person's center of mass are completely unchanged at the transition from one surface to another. This is because the leg stiffness is adjusted such that it completely offsets the difference in surface stiffness.

Insect Running

Before detailed analyses of learning complex motor tasks, Full characterizes the simplest suites of possible "suggestions" from the nervous system that position the mechanical system in state where the behavior can emerge. Full does this by experimentation and modeling. Experiments focus on characterizing the dynamics of animals performing novel maneuvers (i.e. obstacle negotiation, climbing, turning). Three dimensional dynamic whole body models assist in defining the foot and joint conditions used to perform behaviors. Three dimensional dynamic musculoskeletal models link the mechanical system to the control signals (Full & Ahn, 1995).

Full's models suggest that during low intensity terrestrial locomotion, a wide base of support and a low center of mass allows many-legged, sprawled postured animals to be highly statically stable. Animals such as insects and crabs often have at least three legs on the ground. Slow locomotion can be described sufficiently by statics. The dynamics of segments and the body are less important in determining behavior. Classical negative feedback by way of reflexes dominates behavior-based, distributed, neural control found during slow, precise, or novel behaviors. Sensory information from the external or internal environment generates effective local reactions. Continuous feedback control is invaluable when executing novel behaviors and has been shown to be capable of organizing whole behaviors without central pattern generation or feedforward control (?). Neural control by way of feedback gives the animal a degree of versatility and adaptability.

By contrast, effective locomotion at faster speeds can be characterized by a rapid, repetitive, gross behavior. At faster speeds, even many-legged, sprawled postured animals use dynamic gaits. They can operate like inverted pendulums during walking and spring-mass systems during running. Many extraordinarily diverse morphological solutions appear to be adequate for running. Surprisingly, at fast speeds, despite differences in morphology, two-, four-, six- and eight-legged animals produce ground force patterns that are fundamentally similar. All are running or bouncing. The patterns of whole body ground reaction forces produced by the cockroach are similar to those produced by trotting eight-legged crabs, four-legged mammals and running bipeds. In each species, two sets of legs propel the body alternately - a set being one leg for a biped, two for a trotting quadruped, three for cockroaches and four for crabs. The movement of each of these species can be described by a mass atop a spring where the mass represents the body and the virtual leg spring characterizes the behavior of one to four legs. We discovered that each leg of these diverse animals appears to have the same stiffness when force is normalized for body weight and compression for hip height.

In support of the spring-mass hypothesis, the MIT Leg Laboratory designed a dynamic model of a simple abstract insect with telescoping legs attached to a hemispherical body which exhibited dynamic, bouncing locomotion. Three legs acted in a similar way to function as a single virtual leg spring. Although this abstraction captured much of the dynamics of the center of mass of an insect, the model was not representative of an insect's rich morphology. In contrast to most birds and mammals where all legs work in a similar way during hopping, trotting and running, pairs of insect legs each work differently. During a step, the first leg only generates decelerating forces. The force pattern for the second leg appears much like that produced by a human leg, first decelerating and then accelerating the body. The major accelerating force can be attributed to the third leg, which propels the body from behind. In insects each leg operates differently, but three sum to function as one leg of a biped or two legs of a quadruped. Sprawled posture animals not only generate large decelerating forces, but perhaps more importantly, they also produce substantial lateral forces in the horizontal plane that have mostly been ignored.

Two dimensional dynamic model with feedforward control

To understand the importance of morphology for locomotion, Full made a simple two dimensional dynamic model that focuses on the horizontal plane instead of the vertical plane. A rectangle represented the body with the actual mass and the inertia of the insect's body. Massless legs produced force on the body. We input the frequency of leg cycling, the initial position where the feet land, and the pattern and magnitude of individual leg forces. Leg placement and force generation are always relative to the body. The two dimensional insect model moved in a way that was very similar to the animal at its preferred speed. The forward and lateral velocity matched what we measured in the animal. Even the degree of body rotation was similar. To test whether this behavior was robust to perturbations, we changed the momentum of its center of mass as the model ran at a given speed. We speculated that the model could not recover from a velocity perturbation, since it has an extremely simple control system with no explicit feedback circuitry to respond to perturbations. We accelerated it forward to a new, faster, velocity and decelerated it backward to a new, slower, velocity. If we accelerated the model too much, a leg could become fully extended and lose contact with the ground. When this happened, the model stopped as if it fell. Surprisingly, the model converged to the velocity set by the stride frequency over a wide speed range. The horizontal plane model was extremely stable to velocity perturbations of the center of mass. We discovered that at the stable velocity, the net force acting on the center of mass in the middle of the step is purely lateral. This lateral force produces no net change in speed. However, if we decelerated the center of mass to a velocity slower than the stable speed, then the net force acting on the center of mass in the middle of the step was oriented with a forward component. This force produced a net acceleration and a return to the stable speed. The reason for the change is that the net moment on the body at the slower speed under-rotates the body. If we accelerate the center of mass to a velocity faster than the stable speed, then the net force acting on the center of mass in the middle of the step was oriented with a rearward component. This force produced a net deceleration and a return to the stable speed. The reason for the change is that the net moment on the body at the faster speed over-rotates the body. When the velocity is faster than the stable speed, the center of mass moves forward a greater distance while the feet remain in the same position. This produces a large moment arm for the rear leg. The larger moment arm produces a large moment that over-rotates the body. During the next step, when the next tripod lands, a net decelerating force is produced.

Our results suggest that posture can affect dynamic performance in the horizontal plane. To understand the influence of posture on stability, we varied morphology from an upright posture to a more sprawled posture and increased the amount of the lateral force proportionally. Morphology was important for self-stabilizing behavior. Wider stances, and therefore greater lateral forces, resulted in faster recovery by the model to a perturbations. Our two dimensional model tells us that control can reside in the mechanical design of the system and can be simple. The control algorithms are embedded in the form of the animal itself. Control results from the properties of the parts and their morphological arrangement. The mechanical system - the morphology - can determine the extent of self-stabilization. The model suggests that feedforward, simple predictive planning can be effective for rapid, repetitive, gross behavior if it works in concert with the mechanical system.

Use of return maps and the effect of dynamic coupling on control

In perturbing fore-aft velocity of the insect model, Full showed that it was not independent of lateral or rotational velocity. He examined the extent of dynamic coupling in fore-aft, lateral or rotational velocity. Full showed that the recovery rate from perturbations differ considerable. A perturbation in the fore-aft velocity recovers most slowly, lateral velocity at intermediate rates and rotational perturbations are corrected for in nearly one step.

Because of this coupling, Full characterized the stability of all components even if only one is perturbed by using a return map and defining the stable cycle during unperturbed, forward locomotion. Full showed that a given perturbation with correction follows a path that leads to a stability surface and then to a stable point. These simulations have important implications for control. They show how using simple proportional negative feedback actually causes the system to recovery more slowly than if no control was used at all. The natural dynamics of the system tends to simply pull the model away from the stable point in the next step. Knowledge of return maps appear useful, not only for understanding how stability works, but also for integrating passive stabilization with higher order control. The modeling results clearly show that control must work with the natural dynamics, rather than attempting to cancel them out.

Three dimensional dynamic model with distributed feedback control

To examine the role of joints in control and integrate muscle function with whole body dynamics, Full collaborated with Marc Raibert at MIT to design a three dimensional simulation of a running cockroach. The simulation uses the actual morphology (inertia and mass) of cockroach legs reconstructed in our lab, has rigid segments and thirty degrees of freedom. With this three dimensional, dynamic simulation, Full again asked how much control resides in the mechanical system itself, in the morphology. First, Full played in only feedforward torques through the morphology. The simulation became unstable very quickly using feedforward torques alone.

However, the joints of the animal, muscles, tendons and skeletal material operate with complex, non-linear spring and damping properties which come from the force-velocity and force-length functions of muscle as well as the force-length curve of tendon and connective tissue. These properties can offer rapid (zero order) disturbance rejection that operates faster than reflexes. In fact, ? (?) have coined the termed "preflexes" to describe their potential role in stability. Full discovered that playing feedforward torques into their simulation in concert with joint springs and dampers resulted in about five steps of stable forward locomotion. Combining the feedforward component with spring-dampers by using a proportional derivative servo which tracks the animal's joint angle changes results in stable locomotion indefinitely. Preflex contribution is complex and highly non-linear, but appears to allow the dynamic system to be self-stabilizing. No explicit feedback control of global variables such as hopping height, posture, speed was needed for locomotion. The only control was local, at the joints. No communication existed among any of the joints or among the legs.

A model incorporating a feedforward motor pattern with preflexes is not only stable during straight-ahead locomotion, but can also respond to perturbations much like the animal does. To understand the stability of the three-dimensional model, Full introduced brief impulses at the center of mass of a simulated cockroach running at 20 cm/sec. He scaled the impulse to the amount of lateral force generated by the actual animal during locomotion. The simulation showed rapid, self-stabilization to lateral impulses as great as 30 times that found during typical running.

The degree of self-stabilization that results from the morphology of a sprawled postured animal like an insect is surprising. Inherent properties of the musculo-skeletal system can stabilize individual joints. Through mechanical interactions with skeletal segments these musculo-skeletal mechanisms can also serve to stabilize legs, and through interactions with the substratum joint-level control can serve to stabilize the entire body. Feedforward control used in concert with joint preflexes provides simple, but surprisingly effective control of rapid, gross, repetitive tasks. Insects and other arthropods tells us that the performance of legged robots might be enhanced if dynamic bouncing gaits could be attained with sprawled, self-stabilizing postures incorporating rapid, joint-level disturbance rejection.

Insect Flying

Flies (or more properly *Dipterans*), are the group of insects that includes houseflies, horseflies, mosquitoes, gnats, and hoverflies. Flies posess many morphological and physiological specializations that enable them to perform aerial behaviors unmatched by other insects. For example, they are equipped with elaborate gyroscope-like organs called *halteres* which are used for rapid closed-loop control of flight stability (?). In addition, flies possess a unique neural architecture within their eyes that endows them with the fastest known visual systems (Strausfeld & Lee, 1991) (for

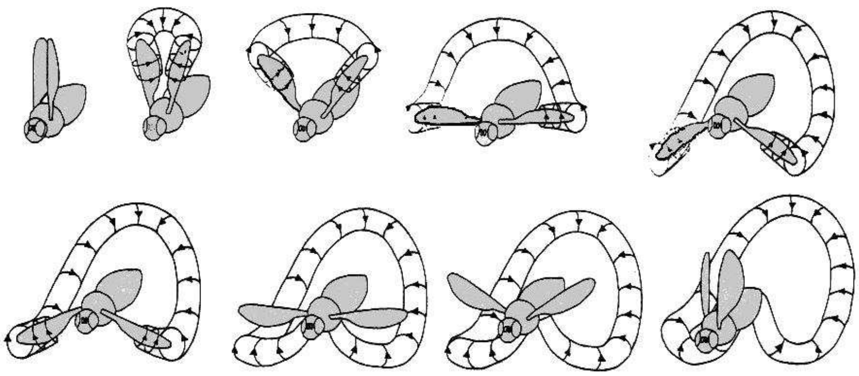


Figure 3: Cartoon reconstruction of vortex loop formation by *Drosophila* from (Dickinson and Götz, 1996). Learning wing interactions with air ("wing gaits") which can efficiently generate lift and thrust will be based on abstracting wing motion and key fluid flow structures.

comparison, the visual system of humans is approximately 10 times slower than that of a typical fly). Unlike most other insects, however, flies possess only a single pair of wings, which makes their aerodynamic mechanisms easier to study.

Flight is a costly form of locomotion, especially for small hovering animals (?). A hoverfly must generate enough mechanical energy to support three terms: aerodynamic power, inertial power, and parasitic power (?, ?). Aerodynamic power is the energy required to overcome drag on the flapping wings and to accelerate air downwards to provide weight support. Inertial power is the cost of accelerating and decelerating the mass of the wing at the beginning and end of each stroke. Parasitic power is the drag on the body caused by its forward motion. Because of the high flapping frequencies employed in flight, the inertial power requirements can represent nearly 50% of potential flight costs (?). The large asynchronous power muscles that fill most of the thorax are specialized to provide the mechanical power for flight (?) (?). By tracking muscle strain with high-speed video, we showed that these muscles operate with a peak-to-peak strain amplitude of about 3-4% in Drosophila (?). One possible explanation for high power output may be high mechanical efficiency (?); (?). Alternatively, insects may recover mechanical energy from one stroke to the next via elastic storage. Respirometric experiments within the flight arena have indicated that the power muscles have a relatively low efficiency of about 10%, but that flies *do* appear to use elastic storage (?); (?), either in the cuticle of the thorax or in the muscle itself (?); (?).

Reverse engineering of insect flight has shown that conventional steady-state aerodynamic mechanisms are not sufficient to keep insects airborne (Ellington, 1995). Instead, insects use the non-steady state mechanism of delayed stall to generate force coefficients 2–3 times greater than could be achieved by conventional fixed-wing based mechanisms (Ellington, v. d. Berg, Willmott, & Thomas, 1996; Dickinson & Götz, 1996). Such elevated performance is possible because insects flap their wings back and forth at a high angle of attack (30–45 degrees). In each half stroke, the elevated angle of attack induces the formation of a leading edge vortex bubble - a mechanism that is analogous to the high lift strategy employed by swept wing aircraft such as the Concorde (Ellington et al., 1996; Dickinson & Götz, 1996). At the end of each stroke the vortex bubble is shed, the wing rotates, and a new bubble forms as the wing translates in the opposite direction (Figure 3). The reciprocating motion of the wing eliminates the necessity of keeping the leading edge bubble absolutely stable over long translation distances (Dickinson, 1996).

Dickinson started investigation of the complex array of steering muscles by focusing on the muscles that insert on the Basalare, a large nail-like sclerite anterior to the wing hinge. He also studied the influence of small phase changes on the mechanics of the first basalar muscle (B1) (?, ?). When oscillated at wingbeat frequency at a strain magnitude comparable to that which occurs in vivo, the muscle always performs negative work. However, the amount of work absorbed and the dynamic stiffness of the muscle are quite sensitive to activation phase. The B1 probably functions as an active elastic element whose stiffness can be altered by the nervous system. By advancing the phase of the B1 within the stroke cycle, the fly increases its stiffness, resulting in an anterior shift in the mean position of the basalar sclerite. Dickinson and his colleagues have further mapped out the connections between the steering muscles and both strain-sensitive sensors and coriolis force sensors through the CNS (?, ?).

Flies are not passively stable. They remain stable while flying only by constantly monitoring their translational and rotational motion and using that information to modulate the production of aerodynamic forces. Two sensory structures are especially important in flight control: compound eyes that are used for visual orientation, and the gyroscope-like halteres that encode angular rotation about the yaw, pitch, and roll axes. The constant output from these sensors is used to modify the activity of flight muscles and to alter aerodynamic forces. These flight control reflexes are fast and robust enough to keep a fly stable in the face of large external perturbations and rapid voluntary maneuvers.

We have recently (?) identified at least eleven distinct control muscles at the base of the haltere in Calliphora, similar to those previously described in Tipula and Tabanus. By comparing the origin and insertion sites we have constructed the tentative homologies with the steering muscles of the wing. At least two of the haltere steering muscles are strongly affected by both visual motion and head movement. Thus, input from descending visual and mechanosensory interneurons might alter haltere kinematics to mimic the inertial forces that occur during corrective reflexes. The haltere-to-wing muscle reflex would respond to this 'virtual' flight perturbation by modifying stroke kinematics, resulting in an active change in flight trajectory.

Insect flight
Our hierarchical model of flight control levels starts first with the resonant mechanical system composed of the fly thorax and wings. The second level includes local strain gauges to measure dynamic wing loading and maintain appropriate beat amplitude and frequency. The third level of control provides rotational stability through the halteres system. The fourth level of control uses visual feedback for orientation control and stabilization.

C.3 A unified approach

We have discussed some promising new approaches to adaptive motor control derived from AI and control theory, and we have sketched our basic findings and some ideas about the structure of the control systems at the action-selection and biomechanical control levels. Based on the results of the experiments already described, we expect to be able to place them in detailed correspondence with appropriately structured adaptive hierarchical controllers, as discussed in Section C.2.1. In order to complete the model of biological behavior as reinforcement learning, we must "reverse engineer" the adaptive behavior to discover what effective reward function is being used. Once this is done, we should be able to predict adaptive behavior in new circumstances and verify those predictions.

C.3.1 Reverse engineering of reward functions

Reinforcement learning has been found to be a powerful method for creating optimal or near-optimal controllers for real tasks (Dietterich, 1997), and it is natural to propose it as a mechanism that might underlie some learning behaviors in natural systems. Indeed, connections are being made currently between reinforcement learning and operant conditioning models using, for example, ACP networks, with some attempts to model actual animal learning (see, e.g., Schmajuk & Zanutto, 1997; Touretzky & Saksida, 1997). There is also some direct evidence from neurophysiological structure that reinforcement learning based on reward signals occurs in bee foraging (Montague, Dayan, Person, & Sejnowski, 1995) and in songbird vocalization (Doya & Sejnowski, 1995).

In most of this prior work, it is generally assumed that the reward function is fixed and known. For example, in experiments on bees it is assumed to be the rate of nectar ingestion: L. and A. (1997) conclude that "the energy expenditure of the foraging bee is controlled by a motivational drive whose intensity depends on the reward rate at the food source," and Montague et al. (1995) cite evidence of a "neuron with widespread projections to odour processing regions of the honeybee brain whose activity represents the reward value of gustatory stimuli."

It is our belief, however, that *in examining animal and behavior we must consider the reward function as as unknown to be ascertained*. The reasons for this belief are straightforward:

- The specification of a given reward function is an empirical hypothesis and may turn out to be wrong. For example, in studying the control of locomotion, some researchers assumed that the objective function being maximized by the controller would be energetic efficiency (?), but this turns out to be a very poor predictor of phenomena such as the speed at which horses switch gaits from trot to canter to gallop (?).

- The reward function is likely to be a complex, nonlinear multiattribute function of the sensory stimuli available, and it seems unlikely that the shape and parameters could be determined by a human expert *a priori*. One might guess that relevant attributes in running might be speed, efficiency, stability against perturbations, wear and tear on muscles, tendons, and bones, etc. The relative weightings and the functional form of the combination surely can only be determined empirically.

In order to model natural control behavior using optimal control ideas, and to model natural learning using reinforcement learning ideas, we must have a reward function for our control and learning algorithms to maximize. To do this, we must first solve the following computational task, which we call *supervised reinforcement learning* [[better name?? reward learning?? inverse reinforcement learning??]]:

Given 1) measurements of a controller's behavior over time, in a variety of circumstances, 2) measurements of the sensory inputs to that controller; 3) a model of the physcial system being controlled.

Determine the reward function that the controller is optimizing.

Given an assumption of optimization, this computational task is well-defined. Notice that is the dual of unsupervised reinforcement learning, where the task is to determine optimal behavior given the reward inputs.

To the extent that behavior is inconsistent with an optimization principle (i.e., no reward function is consistent with all the observed behaviors), we might define computational tasks to determine a "best-fit" reward function under some approximation metric, or to determine restricted sets of circumstances under which various reward functions give good models of the observed behavior. Another version of the problem occurs when observing the *learning* of optimal behaviors—in thi case, the behavior does not maximize the reward function, but observing the adaptation of the behavior over time may suffice to recover the reward signal being used.

To our knowledge, after a search of the literature in several fields, this computational task has not been studied in any generality. The closest work is in economics, where the task of *multiatribute utility assessment* has been studied in depth—that is, how does a person actually combine the various attributes of each available choice when making a decision. The theory is well-developed (Keeney & Raiffa, 1976), and the applications numerous (e.g., Knox & de Chernatony, 1989, discuss utility functions for fizzy water). However, this field studies only *one-shot* decisions where a single action is taken and the outcome is immediate. For historical reasons, the study of optimal *sequential* behavior (from which theories of Markov decision processes and reinforcement learning emerged) occurred largely in operations research and control theory, where the *objective function* is specified by the designer.

The research problems that we shall pursue are several:

- What are efficient algorithms for solving the supervised reinforcement learning problem?
- Under what circumstances can we determine the existence of a consistent reward function? To what extent is the reward function uniquely recoverable?
- What is the computational complexity of supervised reinforcement learning, given assumptions about the functional form of the reward function?
- How much observation is required to determine an estimated reward function that is within ϵ of the true reward function?
- Can we design experiments to recover the reward function from a minimal amount of observation?

In solving these problems, we will draw on methods from mathematical economics (Keeney & Raiffa, 1976), computational learning theory (Kearns & Vazirani, 1994), and adaptive control theory (Sastry & Bodson, 1989).

Considering the design of possible algorithms, the most obvious approach is to use the approach of gradient-based error minimization commonly used in neural networks, together with a parameterized reward function of fixed (but possibly complex) functional form. Thus, we define an error measure $E(B_H(\mathbf{w}), B_T)$, where $B_H(\mathbf{w})$ is the behavior predicted for an optimal controller using a reward function parameterized by $\mathbf{w}$, and B_T is the observed behavior. From this, we can compute $\partial E/\partial \mathbf{w}$. One important question will be how to compute this gradient efficiently; presumably, it can be done in an obvious way by carrying the differential operator through the optimization algorithm for the behavior. More elegant solutions may exist in special cases.

[[Put in intro?— Utility assessment has proved an invaluable tool in modelling and predicting human choice behavior for the one-shot case, and it has had a profound influence on policy decisions ranging from transportation infrastructure planning to medical intervention choices in life-threatening situations. We believe that the generalization of this problem to handle the case of *sequential* behavior will yield very great benefits, both scientific and social, and a range of fascinating computational and experimental research questions.]]

[[The reward function that a controller optimizes may be by far the simplest predictive description of the controller. This is particularly true in the case of learne dbehaviors, where the reward function must be locally and quickly derivable from the immediate sensor inputs.]]

C.3.2 Methodological issues in modelling natural systems

The basic experimental approach for applying supervised reinforcement learning is as follows: 1) examine behaviors under a variety of conditions; 2) recover a reward function; 3) predict behavior under new conditions; 4) compare prediction and experiment.

There are obviously ways in which this might be hard to carry out. First, it is possible that natural controllers animals cannot be described as optimizing any particular combination of attributes; they may just have particular adaptive systems that work differently in different settings and are not *consistent* over all settings. For example, perhaps when running downhill my equilibrium speed effectively equate a 10% speed increase against a 0.001 risk of falling over; but on level, rocky ground my control system actually takes much bigger risks for much smaller speed increases. Empirical evidence that substantial inconsistency exists would be of great scientific interest in its own right.

Second, it may not be easy to identify relevant attributes that are actually measurable by the animal's sensory system. For example, one objection to the model of energy efficiency as an objective function is that animals cannot measure their energy consumption in real time.[6] To the extent that this is true, we share this objection for the case of adaptive controllers—the reward must be available directly from sensor data. This simply limits the space of hypothesized reward functions. For evolved non-adaptive controllers, however, the implicit reward function being optimized need not be sensed, which presentes greater methodological difficulties.

This brings up the general question of the relationship between the proposed research and evolutionary theory. Evolution defines a "fitness function" in terms of reproductive success (but see Dawkins, 1976, for a discussion of the *locus* of optimization). We see the reward function as inherent in the design of the neural system either explicitly, to allow reinforcement learning, or implicitly, if the behavior is already optimzied by evolution. The reward function should be that function whose optimization over time maximizes fitness. Obviously, this is a hypothesis that may be quite false.

In the field of biology, the major swing over the past 10 years has been *away* from assuming that evolution necessarily leads to optimizing anything (Losos & Mikels, 1995).[7] However, the anti-optimization view applies principally to *morphology*, where it is clear that most physical designs retain suboptimal aspects of earlier designs evoled for different purposes—for example, human hands evolved from quadruped forelimbs designed for other purposes. However, given the structure and organization of the physical and control system, there is still an optimization problem to specify the control "algorithm." This is exactly the structured, hierarchical adaptive control problem defined in Section C.2.

C.3.3 Experimental research on natural reward functions

[[describe expts on human running, throwing/juggling, diving. Insect flight and cockroach running may be harder—what happens when you present modified tasks, loads, body modifications etc?]]

Cockroach running

Full's perturbation studies have shown that animals alter their response to these perturbations over time. Animals learn to avoid perturbations and select routes which appear to maximize stability at the highest speed. Animals appear to be updating control laws of the feedforward or feedback systems on each trial. None of these responses have been examined systematically. For the most part, they have been considered noise. Now that Full has hypotheses concerning the control offered by the mechanical system and the potential contribution of local feedback at the joints, testable objective functions can be generated in collaboration with team members. The dynamic model of the cockroach in

[6] Possibly they can estimate mechanical work done, using the spindle organs for sensing muscle forces, but this does not equate directly to energy consumption (Heglund, Fedak, Taylor, & Cavagna, 1982). More recent analyses suggest, however, that a relationship can be ascertained between the two quantities (Kram & Taylor, 1990).

[7] Ironically, the recent upswing in evolutionary psychology operates *against* the view engendered by Kahneman et al. (1982) that the theory of optimal decisions is largely irrelevant to psychology.

conjunction with experimental perturbation test will provide a testbed of hierarchical control models of reinforcement learning.

Human locomotion

[[From CF:
Hypothesis: The movements of the center of mass during running are determined by the passive dynamics of the runner's spring-mass system. The runner can control the movements of their center of mass by adjusting their stance leg stiffness and stance leg touchdown angle in between ground contact phases.

Evidence from biomechanical studies of running suggests that humans strongly prefer to use one touchdown leg angle at a given speed (He et al., 1991; Ferris et al., in press). If they are faced with altered environmental factors at a given running speed, they adjust leg stiffness while keeping leg angle the same. If a person is running at a given speed on a hard surface, there is a single leg stiffness value that can be used with the preferred leg touchdown angle to produce a symmetrical ground contact phase (Fig. 1CTF). If, for example, the leg is too stiff, the stance leg takeoff angle will be smaller than the stance leg touchdown angle (Fig. 2CTF). In addition, the forward velocity of the center of mass will decrease between the beginning and end of the ground contact phase. The opposite trends will be observed if the leg is too compliant.

When a person runs on a compliant elastic surface, it is similar to adding another spring in series with the leg spring. Our findings have shown that if the person is given plenty of time to adapt to the surface stiffness, they always adjust leg stiffness to offset differences in surface stiffness (Ferris et al., in press). The touchdown and takeoff angles of the leg remain the same regardless of surface stiffness. Because of the adjustment to leg stiffness, the center of mass can follow a symmetrical trajectory on all surfaces. Furthermore, many kinematic and kinetic aspects of running remain the same on all surfaces, including ground contact time, stride frequency, peak ground reaction force, and total vertical displacement of the center of mass during the ground contact phase.

If a person suddenly and unexpectedly encounters a change in surface stiffness, it makes sense that they will use the same combination of touchdown leg angle and leg stiffness for their first step on the new surface as they used on the previous surface because they are given no warning that they need to make adjustments. Based on our hypothesis, we predict that the trajectory of the center of mass will not be symmetrical during the first contact phase on the new surface. The asymmetry will occur because the leg stiffness will not be correctly matched to the touchdown leg angle and surface stiffness. For example, if a person is running steadily on a very compliant surface and unexpectedly encounters a hard surface, they will use the leg stiffness appropriate for the very compliant surface during their first contact phase on the hard surface. As a result, their stance leg will be too stiff for the hard surface. If our hypothesis is correct, the movements of the runner's center of mass will follow an asymmetrical trajectory during the first contact phase on the new surface, as dictated by the dynamics of the runner's spring-mass system. The stance leg angle will be smaller at takeoff than it was at touchdown, and the forward velocity of the center of mass will decrease between the beginning and end of the contact phase. Our preliminary observations have shown this pattern, and thus, they appear to support the hypothesis. We have not examined subsequent steps on the new surface but we do know that after several steps, a person adjusts their leg stiffness to the appropriate value and uses the same stance leg touchdown angle on all surfaces.

Our preliminary observations have also revealed that with practice a person can learn to make the leg stiffness adjustment during the aerial phase between the last foot contact on the first surface and the first footfall on the second surface (Ferris et al., in review). In our first study examining sudden transitions from one surface to another, ten practice runs with a specific combination of surfaces were sufficient for a subject to become capable of making the transition perfectly. We believe that human and animal runners have learned the magnitude of leg stiffness adjustment that is required for transitions between common surfaces. For example, a human runner probably can pre-set his/her leg stiffness to approximately the correct value at the transition from concrete to grass based on prior experience.

We have 3 specific aims:

A) To determine if passive dynamics determine the trajectory of the center of mass during the first step after an unexpected change in surface stiffness.

To achieve this aim, we will compare the trajectory of the center of mass in runners to the trajectory predicted by a spring-mass model on a compliant elastic surface. Based on our hypothesis, we predict that they will match. If we are correct, it means that there was not effective neural intervention during one ground contact phase. Rather, the nervous system controls running by adjusting leg stiffness and touchdown leg angle in the time between ground contact phases. We will do these experiments at 2 speeds: a moderate speed (3.0 m/s) and a high speed (5.0 m/s). One might expect that neural intervention during the first ground contact phase would be more likely at the moderate speed (3 m/s) where the ground contact time is 280 ms rather than at the high speed (5 m/s) where the ground contact time is substantially shorter (200 ms). We define "effective neural intervention" as a significant deviation of the trajectory of the runner's center of mass from the trajectory predicted based on a spring-mass system. We have written a Labview (National Instruments) program that allows us to predict the trajectory of the mass in a spring-mass system with any combination leg stiffness, touchdown leg angle, landing velocity, and surface stiffness. It is based on the techniques described in detail by McMahon and Cheng (1990) and Blickhan (1989).

B) To examine how a person adapts between the first unexpected footfall on a new surface and the time several strides later when the person becomes fully adapted and stable.

We predict that a runner rapidly adapts to a new surface by adjusting leg stiffness and leg touchdown angle in the time between footfalls. Because their leg stiffness will not be correct for the first footfall on a surface, they will either accelerate or decelerate. If, for example, they transition from a compliant surface to a hard surface, their stance leg will be too stiff and their forward speed will decrease in the first contact phase. In order to return to their desired speed, they will use a combination of adjustments to leg stiffness and stance leg touchdown angle.

C) To examine the learning process that occurs between the first time that a subject unexpectedly hits a new surface until the subject learns to completely adjust leg stiffness during the aerial phase just before the first footfall on the new surface.

experimenting with different leg stiffnesses during multiple practice runs. Because of the evidence that humans strongly prefer to use a certain value of leg touchdown angle at a given speed, we believe that subjects will not experiment with different touchdown leg angles. In every practice run, they will use the same touchdown leg angle and only alter leg stiffness.

In the experiments, we will make measurements of leg stiffness and touchdown leg angle during the last step on the first surface and the initial step on the second surface to assure that the subjects are not "pre-adjusting" leg stiffness before they arrive at the second surface. We monitored this in our recent study and found that subjects do not pre-adjust leg stiffness (Ferris et al., in review).]]

C.3.4 Insect flight

In deciding on an appropriate performance measure to select for a natural or artificial flying insect, it is important to keep in mind that nature evolved a system of actuators, kinematics, control, and mechanical structure to solve a complex aerodynamic problem. Learning sufficient conditions for flight is perhaps the more interesting challenge rather than learning optimal flight control. For an early insect species with primitive flight capabilities any aerial mobility provides a big competitive advantage, while unstable flight leads to crashing and drowning. Thus, a logical performance measure is “mean time to crash”. We will also examine performance measures that lead to flight with limited actuator power, limited wing degrees of freedom, and small range of motion which will be most appropriate to our MFI.

A higher level learning approach we will examine is *abstracting* the fundamental primitives of fluid flow generated by wing motion. That is, desired net circulation patterns and hence net thrust and lift forces can be learned from existing natural examples.

C.4 Applications to Artificial Systems

[[1.5 pages total, SS coordinates]]

[[include description of available/planned infrastructure)]]

C.4.1 Running in artificial cockroaches

C.4.2 Running and throwing [[RI]] in simulated humans

[[SS, 0.5 pages]]

[[use Bregler tracking (refs), biomech models]] [[as well as bio models, useful for animation/VR etc etc]]

C.4.3 Flying in microelectromechanical winged insects

Insect flight involves complicated interactions between compliant wings beating at frequencies up to several hundred Hertz and moving fluid to generate lift and propulsion forces on the insect body. Analytical models and simulations for unsteady fluid flow are notoriously computationally difficult, and in any case, would provide little guidance for developing effective wing control strategies. While it is unlikely that insects do much individual learning of flying skills, nature has adapted and created control systems for insect flight. Insect flight control presents a grand challenge to engineer adaptive control systems which can learn the essential sensorimotor control loops required for high speed, stable flight.

Our experimental test bed will consist of both real insects and artificial wing systems for which we can measure net flight forces and visualize fluid flow using particle image velocimetry (PIV). (This work is already supported by an ONR MURI grant to co-PIs Dickinson, Fearing, Pister and Sastry). The key learning element we need to understand is the essential relation between wing motion and desired fluid flow structures such as vortices. Learning and copying simple wing motion is not sufficient as our artificial wing systems will not have the degrees of freedom or motion range of the real insect wings. As the wings are beating at 100 Hz, we must consider off-line learning which can analyze several cycles of wing motion and its effect on net forces, and integration of wing control with visual feedback.

We will consider flying insect "learning" in the sense of rapid evolution of flight control strategies. This "evolution" must incorporate basic motor control as well as sensorimotor integration. The following are example learning tasks required for stable and robust insect flight:

- wing motion primitives, e.g. shape parametric learning using 2 DOF basis functions for wing motion
- wing coordination and flight behavior primitives
- learning control: integrating sensorimotor feedback (visual+wing beat/amplitude/phase) with optical flow sensors used for attitude and speed
- learning minimum power flight strategies (optimal wake interactions)
- learning by observation of real insect lift, drag, wing vortices, and net forces

C.5 Roles of the senior investigators

[[SS - update]]

The senior investigators bring to this project a wide range of expertise in analytical and experimental techniques. Russell and Sastry will focus on Artificial Intelligence, control, reinforcement learning, and theoretical analyses of convergence and robustness of learning control techniques. Fearing and Pister will be concerned with the sensing and control of the micromechanical flying insect physical testbed for learning wing and flight control, applying their expertise in micro-robots and MEMs.

The other half of our team brings in the perspective on how biological systems are controlled and can learn new behaviours. Ivry will study high-level control in humans, including sequencing and learning—essentially providing the bridge between the AI models of complex behavior learning and the low-level physical models. Farley and Full provide experimentally determined models of human motor control—phenomena to "explain" in mechanistic terms. They will design and carry out experiments to test new hypotheses about how humans learn proficient behaviours. Dickinson works at the neural level to discover the control mechanisms used by flies and will work with the MFI team to develop strategies for learning optimal wing motion to generate necessary lift forces.

Our team brings together a wide range of backgrounds with a common interest in understanding how complex behaviors in dynamical systems can be learned. Our approach incorporates understanding learning control from the bottom level of fluid and ground interactions in flying and running, up to the highest cognitive levels required in a task such as driving in traffic. Our belief is that there are underlying principles in learning techniques which are best understood and developed by creating a unified model which can address complex behaviors from top to bottom.

C.6 Results from prior NSF support

This section describes the results of prior NSF support to Profs. Fearing, Dickinson, Full, Pister, Russell, and Sastry. The other PIs have not received support from NSF.

C.6.1 Results from prior NSF support—Ronald Fearing

IRI-9157051: PYI Dextrous Robotic Manipulation Total NSF Funding $243,754, 8/1/91 — 7/31/96.

The purpose of this project was to understand how to use tactile sensing for robotic object manipulation and tele-taction. We used a two fingered robot hand to demonstrate closed-loop object reorientation using tactile sensors to provide contact location and contact curvature in real time (Nicolson and Fearing, 1995). We determined noise bounds for distinguishing between smooth curves and edges. (Nicolson and Fearing 1993). We were the first to to use curvature information from a tactile array sensor in closed-loop dextrous manipulation. We developed a tele-taction system using a remote tactile sensor communicating with a 5 by 5 element pressure display which conveyed a remote contact to the user's finger. By proper choice of antialising and reconstruction filters, preliminary results showed human contact localization to better than one fifth of the display element spacing (Cohn, Lam, Fearing 1992). We developed a series of capacitive array sensors ranging in size from an 8 by 8 element sensor in 1 mm square to a 3 by 16 element sensor 25 mm in diameter (Gray and Fearing, 1996). The capacitive technology is stable, reasonably linear, and has adequate sensitivity.

We studied a complicated dextrous manipulation task, threading a nut on a bolt, and developed physically realistic simulations of multiple contacts between the two components as they are inserted (Nicolson and Fearing, 1993a).

We also pursued dextrous manipulation at the milli-scale by developing a credit-card size planar robot system with air bearing supported 2 DOF mobile platforms less than 1 cm in diameter and capacitive position sensing (Fearing 1992a, 1996; Furuichi and Fearing 1996). The idea is to use coordinated motion of multiple miniature platforms to manipulate several small objects in parallel, rather than designing special purpose grippers. We also identified key adhesive force limitations to manipulation of microparts (Fearing 1995), and surveyed actuator limitations for micro-robots (Fearing 1992b).

In studying dynamic legged locomotion, we developed strategies which exploit the non-linear dynamics of a planar Acrobot style mechanism to achieve sliding and hopping gaits with only a single actuated degree of freedom (Berkemeier and Fearing 1992, 1994). Our algorithm uses a set of basis functions and a flight-time constraint to have the robot land in a stablizable posture.

In all, 6 graduate students and 5 undergraduates received support from this project. An additional 1 graduate student and 14 undergraduate students participated in the research supported by the PYI grant.

Availability of Results to Research Community

We have been trying to promote tactile sensing research by sharing the capacitive tactile sensing technology that has been developed in the last few years in our lab. Our schematics and sensor design techniques have been posted on our web page (http://robotics.EECS.Berkeley.EDU), and circuit boards have been made available at cost to 5 robotics labs: Harvard, Stanford, Hong Kong Science University, Univ. of Alberta, and CMU. We have also hosted visitors from Harvard, Hong Kong Science University, Stanford, and Univ. of Alberta in our lab to show them sensor construction techniques. The planar milli-robot technology has been shared with Prof. Isao Shimoyama of Univ. of Tokyo.

Publications

M.D. Berkemeier and R.S. Fearing, "Control of a Two-Link Robot to Achieve Sliding and Hopping Gaits", *IEEE Int. Con. on Robotics and Automation*, Nice France, May, 1992.

M.D. Berkemeier and R.S. Fearing "Control Experiments on an Underactuated Robot with Application to Legged Locomotion", *IEEE Int. Conf. Robotics and Automation,* San Diego, CA May 10-12, 1994, pp. 149-154.

S.R. Burgett, K.S.J. Pister, and R.S. Fearing, "Three Dimensional Structures Made with Microfabricated Hinges", *DSC-Vol. 40 Micromechanical Systems*, ASME Winter Annual Meeting, Anaheim, CA Nov. 8-13, 1992.

M.B. Cohn, M. Lam, and R.S. Fearing, "Tactile Feedback for Teleoperation" SPIE Conf. 1833, Telemanipulator Technology, Boston, MA, Nov. 15-16, 1992.

R.S. Fearing, G. Moy, and E. Tan, "Some Basic Issues in Teletaction", *IEEE Int. Conf. Robotics and Automation* , Albuquerque, NM April 1997.

R.S. Fearing, “A Planar Milli-Robot System on an Air Bearing”, in *Robotics Research the 7th International Symposium,* edited by G. Giralt and G. Hirzinger, pp. 570-581, London: Springer-Verlag 1996.

R.S. Fearing, “Survey of Sticking Effects for Micro-Parts”, *IEEE Int. Conf. Robotics and Intelligent Systems IROS ’95*, Pittsburgh, PA August 7-9, 1995.

R.S. Fearing, “A Miniature Mobile Platform on an Air Bearing”, Third Intern. Symp. on Micro Machine and Human Science, Oct 14-16, 1992a, Nagoya, Japan.

R.S. Fearing, “Micro Structures and Micro Actuators for Implementing Sub-millimeter Robots”, in *Precision Sensors, Actuators and Systems*, edited by H.S. Tzou and T. Fukuda, Kluwer Academic Publishers, Dordrecht The Netherlands, 1992b.

H. Furuichi and R.S. Fearing, “A Planar Capacitive Micro Positioning Sensor”, 7th Int. Symp. on Micro Machine and Human Science, pp. 85-90, Nagoya, Japan, Oct. 2-4, 1996.

B. Gray and R.S. Fearing, “A Surface Micromachined Microtactile Sensor Array”, *IEEE Int. Conf. Robotics and Automation,* Minneapolis, MN April 1996.

E.J. Nicolson and R.S. Fearing, “Compliant Control of Threaded Fastener Insertion” *IEEE Int. Conf. Robotics and Automations*, Atlanta, GA May 3-5, 1993a.

E.J. Nicolson and R.S. Fearing, “Sensing Capabilities of Linear Elastic Cylindrical Fingers”, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems,* Yokohama, Japan, July 1993b.

E.J. Nicolson and R.S. Fearing, “The reliability of curvature estimates from Linear elastic tactile sensors” *IEEE Int. Conf. on Robotics and Automation,* Nagoya, Japan, May 1995.

C.6.2 Results from prior NSF support—Robert Full

DCB 90-58138: Presidential Young Investigator Award , 7/1/90–6/30/95, $ 25,000/year plus matching

Project summary

The present project made possible: 1) the first synthesis of the mechanics, energetics and muscle function in swimming, crawling, running, jumping and flying invertebrates (Full, 1997); 2) the development of a measurement system (muscle lever) to test hypotheses of muscle function under conditions used by animals during locomotion, and 3) the creation of the first fully three dimensional, interactive, dynamic musculo-skeletal model in comparative physiology and biomechanics. Two dynamic models were created - a three dimensional insect leg composed of thousands of polygons from scans of an actual insect leg with all twenty-one muscles and a thirty-six degree of freedom complete body with legs. The models are dynamic, react to gravity and ground reaction forces, have inertia and can become unstable. The development of these models resulted DIRECTLY from the new collaborations with engineers and robotics companies encouraged specifically by the PYI award. Not only have these collaborations lead to discoveries on animal locomotion, but they have allowed us to biologically inspire the design of legged robots. The models completely revised thinking about running in many-legged, sprawled posture animals - from statically, stable, slow walkers where feedback control and efficiency were the focus to dynamically stable, rapid runners where stability is paramount. The models revealed that much of the control of locomotion may be in embedded in the morphology of the animal itself leaving the feedback components of the nervous system for other critical tasks. The choice of the insect model as a simple musculo-skeletal system proved extremely important to tests of muscle function. To complement the models, we developed the capability to measure muscle function directly under conditions used by the insect during running. We discovered by direct, experimental measurement that anatomical extensor muscles can function as energy absorbers during flexion despite being stimulated at the beginning of extension. Muscle function during behavior simply cannot be predicted from the neural stimulation alone, but must be studied in the context of the internal and external environment. Multiple muscle systems can never be understood from the study of isolated muscle, because they are a complex, dynamically coupled, non-linear, and integrated.

Publications resulting from award

Ahn, A. and Full, R.J. Differential leg extensor muscle function in insects during running. J. exp Bio. In preparation.

Full, R. J., Stokes, D. R., Ahn, A. and Josephson, R. K. Energy absorption during running by leg muscles in a cockroach. J. exp Bio. In preparation.

Full, R.J. Invertebrate locomotion. 1997. In: The Handbook of Comparative Physiology. (Ed. W. Dantzler). Oxford University Press. In Press.

Full, R.J. and Ahn, A. 1995. Static forces and moments generated in the insect leg: comparison of a three-dimensional musculoskeletal computer model with experimental measurements. J. exp Bio. 198, 1285-1298.

Anderson, B.D., Shultz J.W. and Jayne B.C. 1995. Axial kinematics and muscle activity during terrestrial locomotion of the centipede Scolopendra heros. J. exp Bio. 198, 1185-1195.

Ahn, A., Eckel, C.M. and Full, R.J. 1995. Extensor and flexor function in exoskeletons: Musculo-skeletal model of an insect leg. Pro. of the Amer. Soc. of Biomech. 19: 239-240.

Ahn, A. and Full, R.J. 1995. Effect of phase, cycle frequency and activation kinetics on power output by an insect leg extensor during oscillatory motion using a three-dimensional computer model. Physiol. Zool. 68: 45.

Full, R.J., Kram R. and Wong, B. 1994. Instantaneous joint power of running roaches. Physiologist 37: A75.

Full. R.J., Kram, R., and Wong, B. 1993. Instantaneous power at the leg joints of running roaches. Amer. Zool. 33: 140A.

Ahn, A. and Full, R.J. 1993. Predicting static joint moment from estimates of muscle moment. Amer. Zool. 33: 140A.

Eckel, C., Ahn, A. and Full, R.J. 1993. Three dimensional simulation of flexor musculo-apodeme function in insects. Amer. Zool. 32: 29A.

Biewener, A. and Full, R.J. 1992. Force platform and kinematic analysis. In: Biomechanics: Structures and Systems A Practical Approach. (ed. A. Biewener). IRL at Oxford University Press. New York. pp 45-73.

Blickhan, R. and Full, R.J. 1992. Mechanical work in terrestrial locomotion. In: Biomechanics: Structures and

Systems A Practical Approach. (ed. A. Biewener). IRL at Oxford University Press. New York. pp. 75-96.

Ahn, A., Full, R. J. and Earls, K. 1992. Partitioning of joint moment among muscles in an insect leg. Am. Zool. 32(5):31A.

Earls, K., Full, R. J., Ahn, A., Eckel, C. and Weinstein, R. B. 1992. Effect on joint moments of varying muscle location using a 3-D, interactive musculo-skeletal model. Am. Zool. 32(5):51A.

Full, R.J., Blickhan, R. 1992. Generality of a spring-mass model in predicting the dynamics of many-legged terrestrial locomotion. Physiologist. 35: 185.

C.6.3 Results from prior NSF support—Michael Dickinson

IBN-9208765: Neural Circuitry and Aerodynamics of Flight Control 7/1/1994–6/30/1997

Unsteady aerodynamic mechanisms in fly flight: Evidence over the past 12 years has accumulated that the *steady-state* performance of insect wings is too low to keep an animal airborne (Dickinson 1996; Ellington 1995; Spedding 1993), fueling the search for *unsteady* mechanisms that could account for the elevated performance of insect wings. To explore such mechanisms, I constructed a dynamically scaled 2-dimensional physical model of a Drosophila wing on which we could directly measure lift and drag forces as it flapped and rotated (Dickinson 1994; Dickinson and Götz 1993). This research led to two important results. First, the mechanism of delayed stall generates forces at the start of rotation that are substantially greater than those generated under steady-state conditions. Second, the rotation of the wing during stroke reversal can greatly influence the generation of forces. Under the right conditions, the vorticity created by rotation of the model wing may be 'captured' during the subsequent translation resulting in lift that is up to 3 times greater than that generated at steady-state (Dickinson 1994). These results were intriguing, because we found that Drosophila actively control the timing of wing rotation during steering maneuvers (Dickinson, Lehmann, and Götz 1993). On the other hand, our studies also indicated that under other conditions stroke reversal can *lower* force production (Dickinson 1996; Dickinson and Götz 1996). In particular, because it spins with an opposite orientation, the circulation shed at the end of the downstroke can greatly attenuate the development of new circulation during the following upstroke.

To study the aerodynamics of insect flight control, we developed a "virtual-reality" flight arena that allows directly correlation of wing stroke kinematics and flight forces. The device is typically configured in "closed-loop" mode, such that the fly itself controls the motion of its surrounding visual world by changing the relative stroke amplitudes of its wings. Using a tiny laser balance, we can simultaneously measure the total flight force and torsional turning force produced by the fly (Dickinson and Lighton 1995; Lehmann and Dickinson 1997). Thus, we have been able to quantify the kinematic changes underlying the control of flight forces in Drosophila (Lehmann and Dickinson 1997; Lehmann and Dickinson 1998).

Muscle efficiency and elastic storage: The large asynchronous power muscles that fill most of the thorax are specialized to provide the mechanical power for flight(Dickinson and Tu 1997) (Dickinson and Tu, 1996). By tracking muscle strain with high-speed video, we showed that these muscles operate with a peak-to-peak strain amplitude of about 3-4% in Drosophila (Chan and Dickinson 1996a). One possible explanation for high power output may be high mechanical efficiency (Casey 1981; Ellington 1991). Alternatively, insects may recover mechanical energy from one stroke to the next via elastic storage. Respirometric experiments within the flight arena have indicated that the power muscles have a relatively low efficiency of about 10%, but that flies *do* appear to use elastic storage (Dickinson and Lighton 1995; Lehmann and Dickinson 1997), either in the cuticle of the thorax or in the muscle itself (Dickinson et al. 1997; Tohtong et al. 1995).

The control of stroke kinematics by the basalar muscles: We started our investigation of the complex array of steering muscles by focusing on the muscles that insert on the Basalare, a large nail-like sclerite anterior to the wing hinge. The first basalar muscle (B1), is unique in that it fires a single action potential each and every wing beat at a precise phase within the stroke cycle (Egelhaaf 1989; Heide 1983). The tonic contraction resulting from this activity pulls the mean position of the basalar apodeme forward (Tu and Dickinson 1996). In the absence of B1 firing, the apodeme is retracted, and the wing can only flutter back and forth within a narrow stroke envelope. When the apodeme is protracted by tonic firing of the B1, the wing swings through its full stroke amplitude. In addition to its tonic role, the B1 also influences stroke kinematics through small changes in its firing phase. During straight flight, the B1 fires within a narrow time window at the start of the downstroke (Tu and Dickinson 1996). During a steering response, however, the B1 on the outside of turn fires 1-2 ms earlier, during the mid-upstroke (Egelhaaf 1989; Heide 1983). These phase advances elicit an increase in stroke amplitude and a promotion of the wing tip during the downstroke. The large B2 muscle has a similar effect on stroke kinematics, but it does not fire tonically and is active only during steering maneuvers. We adopted the work-loop approach of Josephson (1985) to study the influence of small phase changes on the mechanics of the B1 (Tu and Dickinson 1994). When oscillated at wingbeat frequency at a strain magnitude comparable to that which occurs in vivo, the muscle always performs negative work. However, the amount of work absorbed and the dynamic stiffness of the muscle are quite sensitive to activation phase. The B1 probably functions as an active elastic element whose stiffness can be altered by the nervous system. By advancing the phase of the B1 within the stroke cycle, the fly increases its stiffness, resulting in an anterior shift in the mean position of the basalar sclerite.

Neural control of B1 muscle: The biomechanical studies described above indicate that the firing phase of the B1 is important in determining stroke kinematics. Ablation experiments by Heide (1983) suggested that the synchronization of the B1 motor neuron (MNB1) within the wingbeat is due, not to the input of a central pattern generator, but rather the phasic drive from mechanosensors at the base of the wings and halteres. The halteres are small club-shaped hind wings characteristic of flies that have been transformed through evolution into equilibrium organs; they beat antiphase to the wings during flight and are most sensitive to the rotational Coriolis forces. There are 5 sensory fields at the base of the haltere, comprising roughly four hundred strain-sensitive campaniform sensilla. By filling small groups of neurons in each field, we determined that the central projections of the campaniform cells are highly structured, such that cells originating from different fields project to stereotypic locations within the CNS (Chan and Dickinson 1996b). We also developed a preparation to study the synapse between the haltere afferents and MNB1 (Fayyazuddin and Dickinson 1996). Mechanical oscillation of the haltere within its beating plane elicits a large compound EPSP in MNB1. By ablating individual campaniform fields at the base of the haltere, we traced the source of MNB1's input to dF2. Ion exchange experiments demonstrated that the connection between the dF2 campaniforms and MNB1 is monosynaptic and composed of electrical and chemical components. As with the haltere, electrical stimulation of the wing nerve is also capable of driving MNB1 past threshold.

The motor control system of the haltere: A novel mechanism for the visual control of flight: We have recently (Chan, Prete, and Dickinson 1998) identified at least eleven distinct control muscles at the base of the haltere in Calliphora, similar to those previously described in Tipula (Mickoleit 1962) and Tabanus (Bonhag 1949). By comparing the origin and insertion sites we have constructed the tentative homologies with the steering muscles of the wing. At least two of the haltere steering muscles are strongly affected by both visual motion and head movement. Thus, input from descending visual and mechanosensory interneurons might alter haltere kinematics to mimic the inertial forces that occur during corrective reflexes. The haltere-to-wing muscle reflex would respond to this 'virtual' flight perturbation by modifying stroke kinematics, resulting in an active change in flight trajectory.

Publications

Chan, W P, and M H Dickinson. 1996a. In vivo length oscillations of indirect flight muscles in thr fruit fly Drosophila virilis. J. Exp. Biol. 199: 2767-2774.

Chan, W P, and M H Dickinson. 1996b. Position-specific central projections of mechanosensory neurons on the haltere of the blow fly, Calliphora vicina.. J. Comp. Neurol. 369:405-418.

Chan, W P, F Prete, and M Dickinson. 1998. Visual input to the efferent control system of a fly's 'gyroscope'. Science 289:289-292.

Dickinson, MH. 1996. Unsteady mechanisms of force generation in aquatic and aerial locomotion. Amer. Zool. 36:537-554.

Dickinson, Michael H. 1994. The effects of wing rotation on unsteady aerodynamic performance at low Reynolds numbers. J. Exp. Biol. 192:179-206.

Dickinson, Michael H, and Karl G Götz. 1993. Unsteady aerodynamic performance of model wings at low Reynolds numbers. J. Exp. Biol. 174:45-64.

Dickinson, M. H., and K. G. Götz. 1996. The wake dynamics and flight forces of the fruit fly, Drosophila melanogaster. J. Exp. Biol. in press (199):2085-2104.

Dickinson, M H, C J Hyatt, F O- Lehmann, J R Moore, M C Reedy, A Simcox, R Tohotong, J O Vigoreaux, H Yamashita, and D W Maughan. 1997. Phosphorylation-dependent power output of transgenic flies: an integrated study. Biophys. J. 73:3122-3134.

Dickinson, Michael H, Fritz-Olaf Lehmann, and Karl G Götz. 1993. The active control of wing rotation by Drosophila. J. Exp. Biol. 182:173-189.

Dickinson, M. H., and J. R. B. Lighton. 1995. Muscle efficiency and elastic storage in the flight motor of Drosophila. Science 128:87-89.

Dickinson, M. H., and M. S. Tu. 1997. The function of Dipteran flight muscle. Comp. Biochem. Physiol. A 116A:223-238.

Fayyazuddin, A, and M H Dickinson. 1996. Haltere afferents provide direct, electrotonic input to a steering motor neuron of the blowfly, Calliphora. J. Neurosci. 16:5225-5232.

Lehmann, F- O, and M H Dickinson. 1997. The changes in power requirements and muscle efficiency during elevated force production in the fruit fly, Drosophila melanogaster. J. Exp. Biol. 200:1133-1143.

Lehmann, F- O, and M H Dickinson. 1998. The control of wing kinematics and flight forces in fruit flies (Drosophila spp.). J. Exp. Biol. 201:385-401.

Tu, Michael S, and Michael H Dickinson. 1994. Modulation of negative work output from a steering muscle of the blowfly Calliphora vicina. J. Exp. Biol. 192:207-224.

Tu, M. S., and M. H. Dickinson. 1996. The control of wing kinematics by two steering muscles of the blowfly, Calliphora vicina. J. Comp. Physiol. A 178:813-830.

C.6.4 Results from prior NSF support—Kris Pister

IRI-9321718: Micro Robot Manipulators Total NSF funding $266,000, July 1, 1994 - June 30, 1998.

The impact of the research funded under this award was substantial and broad.

A complete library of components for the fabrication of articulated micro robot arms has been developed. This library is available on the web, and includes the first microfabricated three degree of freedom articulated arm. In the final months of the project we hope to demonstrate the first chip which can crawl under it's own power.

A completely new micromachining process was introducted to the MEMS community (xenon difluoride sacrificial layer etching). This process is now in widespread use around the world, and is available commercially from a major semiconductor equipment vendor.

The use of standard CMOS to make three dimensional structures with integrated sensors and actuators was demonstrated. This has led to more than a dozen publications and one patent, now licenced and being commerciallized by Maxim, a silicon valley analog circuit company.

The first integration of hinged polysilicon MEMS devices and semiconductor laser diodes was demonstrated in collaboration with Prof. Ming Wu. Together with work at UCB and AFIT, this launched a new area in micro optics which is now flourishing worldwide.

Selected relevent publications

G. Lin, K.S.J. Pister, and K.P. Roos, "Novel Microelectromechanical System Force Transducer to Quantify Contractile Characteristics from Isolated Cardiac Muscle Cells", *J. Electrochemical Society*, March 1995, Vol. 142, No. 3, L31-3.

L.Y. Lin, S.S. Lee, K.S.J. Pister, and M.C. Wu, "Three-dimensional Micro-Fresnel Optical Elements Fabricated by Micromachining Technique", *Electronics Letters*, Vol. 30 No. 5, pp. 448-449.

Gunawan, D.S.; Lin, L.-Y.; Pister, K.S.J., "Micromachined corner cube reflectors as a communication link.", *Sensors and Actuators A (Physical)*, March-April 1995, Vol.A47, (no.1-3):580-3.

Richard Yeh, Ezekiel J.J. Kruglick, Kristofer S.J. Pister, "Surface Micromachined Components for Articulated Microrobots", *IEEE/ASME J. MEMS*, Vol. 5, N. 1, March 1996, pp. 10-17.

Patrick B. Chu, Phyllis R. Nelson, Mark L. Tachiki, and Kristofer S. J. Pister, "Dynamics of polysilicon parallel-plate electrostatic actuators", *Sensors and Actuators A (Physical)*, vol.A52, (no.1-3):216-20, March-April 1996.

Richard Yeh, Ezekiel J. Kruglick, Michael E. Klitzke, and Kristofer S.J. Pister, "Towards an Articulated Silicon Microrobot", *ASME Winter Annual Meeting*, Chicago, Nov. 6-11, 1994, DSC-Vol 5502, pp. 747-754.

E. Hoffman, B. Warneke, E. Kruglick, J. Weigold, K.S.J. Pister, "3D Structures with Piezoresistive Sensors in Standard CMOS", *IEEE Micro Electro Mechanical Systems Workshop*, Amsterdam, The Netherlands, 1995, p. 288-293.

Richard Yeh, Ezekiel J.J. Kruglick, K.S.J. Pister, "Microelectromechanical Components for Articulated Microrobots", *IEEE Intl. Conf. on Solid State Sensors and Actuators (Transducers '95)*, Stockholm, June 25-29, 1995, pp. 346-349.

Brett Warneke, Eric Hoffman, and Kristofer S. J. Pister, "Monolithic Multiple Axis Accelerometer Design in Standard CMOS", *Proc. SPIE Conf. Micromachined Devices and Components*, Austin, TX, October 1995, pp. 95-102.

Beverley Eyre, Kristofer S. J. Pister, and Walter Gekelman, "Multi-axis Micro-coil Sensors in Standard CMOS", *Proc. SPIE Conf. Micromachined Devices and Components*, Austin, TX, October 1995, pp. 183-191.

Richard Yeh and Kristofer S. J. Pister, "Measurement of Static Friction in Mechanical Couplings of Articulated Microrobots", *Proc. SPIE Conf. Micromachined Devices and Components*, Austin, TX, October 1995, pp. 40-40.

G. Lin, K. S. J. Pister, K. P. Roos, "Standard CMOS Piezoresistive Sensor to Quantify Heart Cell Contractile Forces", *Proc. IEEE Micro Electro Mechanical Systems Workshop*, San Diego, California, February 11 - 15, 1996, pp. 150-155.

K. Pister, "Smart Dust", SCCAVS Symposium, Micromachining Workshop III, Sept. 25, 1996, Anaheim, CA.

P.B. Chu, N.R. Lo, E.C. Berg, K.S.J. Pister, "Optical Communication Using Micro Corner Cube Reflectors", *Proc. IEEE MEMS Workshop*, Nagoya, Japan, Jan. 26-30, 1997, pp. 350-355.

P.B. Chu, J.T. Chen, R. Yeh, G. Lin, J.C.P. Huang, B. Warneke, K.S.J. Pister, "Controlled pulse etching with Xenon Difluoride", *IEEE Intl. Conf. on Solid State Sensors and Actuators (Transducers '97)*, Chicago, IL, June 16-19, 1997, pp. 665-668.

C.6.5 Results from prior NSF support—Stuart Russell

IRI-9058427: PYI: Architectures and Algorithms for Autonomous Intelligent Systems Total NSF funding \$300,000, 8/1/90–7/31/95. This project covered topics in artificial intelligence relating to learning complex models, learning behaviours, decision making under uncertainty, real-time decision making, and foundations of limited rationality. Results relevant to the current proposal include the following:

- *Expressive models and online state estimation for complex stochastic systems.*
  Several contributions were made to the development of *dynamic Bayesian networks* or DBNs (Nicholson, 1992)—factored-state models of partially observable Markov processes that generalize hidden Markov models and Kalman filters. DBNs allow the representation of very large systems with arbitrary dynamics; we applied them successfully to the problem of extracting high-level behavioural descriptions from video monitoring of freeway traffic (Koller et al, 1994). For any large, partially observable Markov chain, exact state estimation is intractable and standard Monte Carlo methods fail miserably. We developed a new stochastic simulation method (Kanazawa et al., 1995) whose approximation error remains bounded over infinite sequences.

- *Decision making with limited computational resources.*
  The vast majority of real-world decision problems are intractable; limited computational resources must be allocated effectively to generate the best possible decisions. The method of *rational metareasoning* was developed for efficiently estimating the value of a computation step prior to its execution, and hence controlling the computational process. The method proved capable of generating excellent real-time computational behaviour in combinatorial problems, games, and large Markov decision problems (Russell & Wefald, 1991; Tash & Russell, 1994). The approach was extended to provide a rigorous foundation for limited rationality as *bounded optimality* (Russell & Subramanian, 1995), and applied to the construction of optimal composite real-time systems (Zilberstein & Russell, 1996).

- *Embedded system architectures for planning under uncertainty.*
  Using dynamic Bayesian networks as a substrate, we developed a generic architecture for embedded decision systems receiving a continuous stream of noisy data. We combined rational metareasoning with hybrid (reactive and deliberative) decision algorithms to obtain effective behaviour in areas ranging from autonomous submarines (Ogasawara & Russell, 1993) to freeway driving (Forbes et al., 1995).

- *Learning complex probabilistic models from data and prior knowledge.*
  The first effective methods were developed for learning Bayesian networks from incomplete data, including learning DBN models of partially observable Markov processes (Russell et al., 1995).

- *Artificial intelligence as rational agent design.*
  A framework was developed within which most past and current research in artificial intelligence could be unified. This led to the publication of *Artificial Intelligence: A Modern Approach* (Russell & Norvig, 1995), which is the text for more than 65% of AI courses worldwide.

In all, two postdoctoral students, eleven PhD students, five MS students, and seven undergraduate students worked on the project.

Selected Relevant Publications (out of 67 papers total)

1. J. Forbes, T. Huang, K. Kanazawa, and S. Russell, "Towards a Bayesian Automated Taxi." In *IJCAI-95*.
2. K. Kanazawa, D. Koller, S. Russell, "Stochastic simulation algorithms for DBNs." In *UAI-95*.
3. D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards Robust Automatic Traffic Scene Analysis in Real-Time." In *ICPR-94*. *Best Industry-Related Paper Award (Honorable Mention).*
4. G. Ogasawara and S. Russell "Planning Using Multiple Execution Architectures." In *IJCAI-93*.
5. S. Russell, J. Binder, D. Koller, "Local learning in probabilistic networks with hidden variables." In *IJCAI-95*.
6. S. Russell and P. Norvig *Artificial Intelligence: A Modern Approach*. Prentice Hall, 1995.
7. S. Russell and D. Subramanian "Provably bounded optimal agents." *J. Artificial Intelligence Research*, **2**, 1995.
8. S. Russell and E. Wefald *Do the Right Thing: Studies in Limited Rationality*. MIT Press, 1991.
9. J. Tash and S. Russell "Control strategies for a stochastic planner." In *AAAI-94*.
10. S. Zilberstein and S. Russell "Optimal composition of real-time systems." *Artificial Intelligence*, **82**, 1996.

IRI-9634215: Learning complex probabilistic models from data Total NSF funding $380,000, 1/1/97–12/31/99. This project extends one of the lines of work initiated in the previous project. It focuses on extending the expressiveness of probabilistic models, extending the scope of learning algorithms, and developing substantial scientific applications. In the first year, we have obtained the following results:

- *Methods for automatic creation and modification of model structures.*
  A new algorithm, Structural EM (Friedman, 1997) represents a significant development in computational statistical methods. SEM generalizes the well-known EM to incorporate structural as well as parametric learning, while retaining the convergence guarantees of EM. We have applied SEM to learn substantial models, including DBN models of speech generation that substantially outperform hidden markov models for recognition (Zweig & Russell, 1998).
- *Theoretical analysis of sample and complexity bounds for PN and DPN learning.*
  A thorough analysis of sample complexity for learning Bayesian networks appears in Dasgupta (1997), developing tight bounds using covering numbers.
- *Methods for combined model learning and reinforcement learning in partially observable environments.*
  Andre et al. (1997) have shown how DBN models can be combined with reinforcement learning, providing a powerful method for adaptive control of Markov processes.
- *Automated extraction of human driver models from videotapes.*
  We have developed simple DBN models of human drivers and and trained them directly from vehicle tracking data (Oza & Russell, submitted).
- *Adaptive hierarchical control for large Markov processes.*
  For very large Markov processes, tractable control policies must be hierarchically structured. In (Parr & Russell, 1997), a language is proposed for describing partially specified hierarchical policies. Algorithms are given for efficient online learning of optimal policies consistent with the prior specifications. These results may make practical a theoretically rigorous approach to the control of very large systems.
- *Object identification under uncertainty.*
  When observing over time a system consisting of multiple objects, state estimation and model learning require computing the probability that one observed object is in fact the same as another. We developed appropriate probabilistic models and inference algorithms for this problem, and applied them to estimate freeway travel times using video data streams from widely separated camera sites. The resulting paper (Huang & Russell, 1998) received a Distinguished Paper Award at IJCAI 97 (out of 816 submissions).

Selected Relevant Publications (out of 20 papers total)

1. D. Andre, N. Friedman, and R. Parr, "Generalized Prioritized Sweeping." In *NIPS '97*.
2. J. Binder, D. Koller, S. Russell, K. Kanazawa, "Adaptive Probabilistic Networks with Hidden variables." *Machine Learning*, **29**, 213–244, 1997a.
3. S. Dasgupta, "The sample complexity of learning Bayesian networks." *Machine Learning*, **29**, 165–180, 1997.
4. N. Friedman, "Learning belief networks in the presence of missing values and hidden variables." In *ICML-97*.
5. T. Huang and S. Russell, "Object Identification in a Bayesian Context." *Artificial Intelligence*, to appear.
6. R. Parr and S. Russell, "Reinforcement Learning with Hierarchies of Machines." In *NIPS '97*.
7. G. Zweig and S. Russell. "Speech Recognition with Dynamic Bayesian Networks." In *AAAI-98*.

C.6.6 Results from prior NSF support—Shankar Sastry

IRI-95-31837: Millirobotics for tissue manipulation in minimally invasive telesurgery Shankar Sastry, ROn Fearing, Frank Tendick (UCSF), $ 427,355 from August 1, 1996 to 31st July, 1998

Research in the first 21 months has focused on issues of teleoperation and teletaction. A new NSF grant CDA 9726362 starting Sept. 1997, entitled "Virtual Environments for Telesurgery" with co-PIs Barsky and Tendick (UCSF) has been in place for 9 months and aims at developing virtual reality tools for modeling of organs, tissue for training and implementing telesurgery as well as conventional surgery.

There is a group of about 6 students and 2 post doctoral scholars supported on these grants. They will support the KDI initiative in that they provide excellent test beds for skill acquisition and learning of new motor skills for surgical purposes. This is yet another test bed not explicitly mentioned in the proposal.

Teleoperation: We have been studying major issues in teleoperation in the context of telesurgery, including the fidelity-stability trade-off and human factors. In particular, one way to increase safety and performance in telesurgery with time delay is to use hierarchical control. In a hierarchical scheme, the commands of the surgeon are transmitted to a local low-level controller which guarantees safe execution of the commands. Using hybrid control design techniques, we have developed a control law that serves as a low level controller to limit interaction forces, guaranteeing safety under specified disturbances (Cavusoglu et al., 1997). The main principle is to find the largest safe set of velocities and interaction forces, and then to synthesize a control action which will be activated at the boundary of this safe set to make it invariant. As the control action is specified only for the boundary of the set, it will be the least-restrictive controller guaranteeing safety. Effectiveness of this scheme has been shown so far in simulation.

Another major issue in teleoperation is the correspondence between visual and control geometries. In minimally invasive surgery and telesurgery, the surgeon has only a video image of the surgical workspace. We have measured the trajectories of movements performed by subjects under videoscopic conditions, in a virtual environment, and in free space under direct vision. Under all conditions, when moving from point to point, both experienced surgeons and novices make initial coarse movements of 600 milli-second duration, followed by a slow fine phase to intercept the target. The accuracy of the coarse phase determines the total time and accuracy of the movement. Because the coarse phase is fast enough that visual feedback is of little use, this phase must be largely "pre-programmed" by the human visual and motor systems. Distortion of the display and viewing geometry in videoscopic and virtual conditions produces inaccurate coarse phases. We are attempting to develop models and schemes to design displays and teleoperators that minimize geometric distortions.

Teletaction: In the past year, an approximate analytical model of a complete teletaction system has been developed (Fearing et al., 1997). With this model, we can analyze different versions of the tactile display without physically building it. Parameters of interest include elastic layer thickness (used as a reconstruction filter) and stimulator density. Most tactile stimulators can only apply a normal force to the tip of the finger. We are looking at the error caused by the stimulator elements not being able to exert tangential forces. Our model predicts that we can not achieve a sensed signal-to-noise ratio better than 16 dB without presenting surface shear information to the finger. We are conducting psychophysics experiments to determine what effect surface shear has on perception of contact localization, resolution and shape. We have also used psychophysics to measure hysteresis in shape perception (Singh, 1997); we found a wide variation between subjects of a “tactile afterimage”. This afterimage limits perception in tasks such as palpation and lowers display temporal bandwidth requirements.

Technology development has continued for tactile sensors and displays. Two new sensors were constructed: a 4 by 8 array which fits on a 10 mm diameter laparascopic instrument, and a 5 by 5 array which is designed to fit on an 8 mm by 8 mm parallel jaw laparascopic gripper. Our current prototype uses a piezoelectric cantilever and can run at 50 PSI and should give us close to 100 Hz bandwidth (measurements in progress). The active element in the valve is only 15 mm by 3 mm, so potentially we can build a compact valve array.

Telesurgical Workstation: Construction and Testing: We have continued with the construction and testing of a telesurgical workstation with a holding robot and endoscopic robots. The development of the surgical workstation is supported by other grants, but the development of the control laws and the frequency response measurements as well as the modeling and analysis of the actuators is performed on this grant. Two papers giving a survey of our work in this area are (Sastry et al., 1997; Tendick et al., 1997). A recent M. S. thesis also describes the current system measurents (Cavusoglu, 1997).

Virtual Environments for Telesurgery: We have developed a system which is being tested at the medical school for cholsystectomy. Papers on this have been presented at all major medical VR conferences (Cavusoglu et al., 1998; Tendick et al., 1998; Cavusoglu and Tendick, 1997), and the material has been given high-profile coverage in the print and TV media.

1. M. C. Cavusoglu, "Control of a Telesurgical Workstation", *M. S. Project Report*, Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, May 1997.
2. M.C. Cavusoglu, J. Yan, S. Sastry, " Hybrid System Approach to Contact Stability and Force Control in Robotic Manipulators," Proceedings of *12th IEEE International Symposium on Intelligent Control,* Antalya, Turkey, July 1997, pp. 143-148.
3. R.S. Fearing, G. Moy and E. Tan, "Some Basic Isues in Teletaction", *Proceedingg of the IEEE International Conference on Robotics and Automation'*, Albuquerque, NM April, 1997.
4. U. Singh, "Temporal Characteristics of the Human Finger", *M. S. Project Report*, Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley, May 1997.
5. F. Tendick, S. Sastry, R. Fearing, and M. Cohn, "Applications of Micro-Mechatronics in Minimally Invasive Surgery," accepted to appear in *IEEE/ASME Transactions on Mechatronics*, February 1997.
6. S. Sastry, M. Cohn, and F. Tendick, "Millirobotics for Remote, Minimally Invasive Surgery," to appear in *Journal of Robotics Research*, 1997.
7. M. Cavusoglu, M. Cohn, F. Tendick, and S. Sastry, "Laparoscopic Telesurgical Workstation," Proceedings of the *SPIE International Symposium on Biological Optics (BIOS'98)*, San Jose, Jan 1998.
8. F. Tendick, M. S. Downes, M. C. Cavusoglu, and L. W. Way, "Development of Virtual Environments for Training Skills and Reducing Errors in Laparoscopic Surgery" In Proceedings of the *SPIE International Symposium on Biological Optics (BIOS'98)* , San Jose, CA, January 24-30, 1998.
9. F. Tendick, and M. C. Cavusoglu. "Human Machine Interfaces for Minimally Invasive Surgery," In Proceedings of the *19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (EMBS'97), Chicago, IL, October 30- November 2, 1997, pp. 2771-2776.

IRI-9014490: Nonholonomic Motion Planning for Robots Shankar Sastry, $ 319,595, from 8/15/1990 to 1/31/1995.

The aim of this project was the development of methods for planning of trajectories for robotic systems with non-holonomic constraints (that is constraints that can not be integrated, such as rolling without slipping or conservation of angular momentum). Such systems include, for example, Mobile Robots with N trailers, Mobile Robots with N trailers and multiply many steered (such as a fire-truck), Space Robots, with conservation of angular momentum, Multifingered hands rolling on the surface of a grasped object, Satellites and Aircraft, and Submarines. The progress made on the grant resulted in the development of new tools bridging control theory, computational geometry and motion planning for robots and gave solutions to important previously unsolved problems, such as:

1. Parking a mobilerobot with N trailers in a field of obstacles.
2. Reorienting a satellite with one or two rotors
3. Reorienting a robot in space using internal motions of the object
4. Riding a robotic bicycle

In addition to algorithms for steering such system, we developed software tools for animating the schemes and for interactively steering the systems considered. The unusually productive grant resulted in 5 completed doctoral dissertations, several Masters and three Honors Bachelors dissertations.

C.7 Plans for dissemination, institutional commitment

[[NSF: An additional section, up to 2 pages, that describes and justifies specifically a) the plans for dissemination of the results and b) the institutional commitment as to space and equipment.]]

C.7.1 Plans for dissemination of research results

[[SS:1 page]]

[[Anyone for a textbook? We have a good track record...]]

[[From 1997 LIS:
Research Seminars: we will run a bi-weekly interdisciplinary research seminar on *Adaptive Control and Sensorimotor Learning* including control theory, biological, neural, and cognitive modelling, robustness and verification, reinforcement learning, hierarchical control, hybrid systems, vision-based control.]]

[[From Full: The PolyPEDAL laboratory will disseminate results by in an effort to make the most significant interdisciplinary effort possible. They will make a concerted effort to present papers at meetings not typically attended. Biologists will present papers at engineering meetings (such as the invitation we received to the Yale Workshop on Adaptive Controls). Biologists will invite engineers to biological meetings such as Neural Control of Movement, Neurosciences, and Neuroethology.

The PolyPEDAL laboratory will place all published abstracts and papers on the their WWW site (http://polypedal.berkeley.edu). They will provide summaries of the findings by the whole group so that biologists may view and understand data and approaches that they would have not otherwise have seen.]]

C.7.2 Institutional commitment of space and equipment

[[ALL: want any new toys not funded under the MURI grant?]]

The proposed research will take advantage of a very substantial commitment of space and research infrastructure spread over five laboratories. Since this proposal does not include the design and fabrication of physical prototypes (these are covered under other grants) we are not requesting substantial additional capital equipment with the exception of one force platform for biomechanical experiments.

EECS Robotics Lab
The main research projects in the EECS Robotics Lab are currently 1) milli-robotics for surgery and assembly and 2) tele-robotics including force (kinesthetic) and tactile (cutaneous) feedback, and 3) non-linear control, particularly of hybrid systems. General lab infrastructure includes a network of a dozen high performance workstations (Sun and Silicon Graphics), and supporting software such as Matlab, AutoCAD, etc.

The Robotics Lab has facilities for development and testing of MFI systems, particularly the electronic interface and control systems prototypes. For low-precision conventional machining we have a table top CNC mill and lathe. We also have Pentium Pro PCs with data acquisition hardware to use for control of the micro-robot prototypes. (Eventually we will use dedicated controllers such as DSP chips).

Biomimetic Millisystem Laboratory
The Biomimetic Millisystem Laboratory is being created in summer 1998 for design, fabrication, and testing of millimeter and scale mobile systems based on biomimetic principles. The lab, to be used on a shared basis by Engineering and Integrated Biology faculty and students, will contain specialized equipment suited for millimeter scale fabrication and testing, particularly for non-silicon and non-metallic fabrication. The lab creation is being funded by $195K of campus funds with $200K of equipment funding through an ONR MURI.

For rapid prototyping of compliant actuators, sensors, and structures, it will contain a New Wave Research Laser Micromachining System. This machine is capable of cutting holes in polysilicon, aluminum, polyimide with 2 micron resolution. One of the primary uses will be for polymer wing cutting, and etch anchor cutting. We will also be acquiring a particle image velocimetry (PIV) system for measuring fluid flow from model and real insect wings. The lab temperature and relative humidity will be controlled to avoid problems with dimensional changes and surface tension effects, and equipment overheating. We are also including appropriate sink and fumehood combinations for safe handling of any volatile polymers. By use of glass floor-to-ceiling partitions, we will isolate potentially hazardous laser micromachining and imaging (PIV) operations from millisystem assembly and testing areas. A microscope probe station (Micromanipulator Co.) will be used for assembly, inspection and testing of micro-robots.

Berkeley Microfabrication Lab
The Berkeley Microfabrication Lab, located directly above the Biomimetics Millisystems Lab and EECS Robotics Lab, will be used extensively for MEMS fabrication steps for the MFI.

PolyPEDAL Lab
The PolyPEDAL Laboratory studies the locomotion of small, many–legged animals. Our 1800 ft sq laboratory contains four types of work-stations. The Energetics work-stations consist of miniature treadmills, O_2 and CO_2 analyzers for the estimation of the metabolic cost of locomotion. The Mechanics work-stations include high-speed video cameras for three dimensional kinematic analyses and several force sensitive platforms for three dimensional kinetics. The Muscle work-stations consist of computer controlled muscle levers and EMG data acquisition. The Modeling work-stations consist of SGI computers with muscle and whole body dynamic modeling programs. We also have access to a biological visualization center and an ecological scanning electronic microscope which allows pictures of living tissues.

UC Berkeley Locomotion Lab
The UC Berkeley Locomotion Laboratory has facilities and equipment for the study of the biomechanics and energetics of vertebrate legged locomotion. The laboratory (3900 sq.ft.) includes an overground locomotion site with a runway, force platforms, high speed video system, and telemetered EMG system for studying overground locomotion in humans and other large animals. The laboratory also includes a "force treadmill", a device for measuring the vertical, horizontal, and lateral ground reaction forces during human locomotion on a treadmill; this device allows quantification of the dynamics of many consecutive strides of locomotion. In conjunction with the force treadmill, there is a metabolic workstation for measuring the rate of energy consumption of humans during locomotion. Finally, the Locomotion Laboratory has a workstation for studying the kinematics and mechanics of locomotion of small vertebrate animals and the mechanical properties of their muscles. These include a miniature force platform, an EMG system, and a muscle mechanics workstation.

Dickinson Lab: Virtual-reality insect flight arena
This unique facility allows simultaneous measurement of wing stroke kinematics by high-speed video cameras and flight forces (both total flight force and torsional turning force) using a tiny laser balance. On the inside of the arena are projected virtual-reality images in order to measure the effects of sensory input on insect flight control. The device is typically configured in "closed-loop" mode, such that the fly itself controls the motion of its surrounding visual world by changing the relative stroke amplitudes of its wings.

C.8 Performance goals

[[SS,1 page]]

[[NSF: For projects with a three year budget exceeding $1.0 million (optional for projects with smaller budgets), an additional page describing realistic performance goals for each year.]]

[[From Full:

Year 1: Characterize the variability of individual insects to perturbations during running over irregular terrain. Measure three dimensional kinematics and ground reaction forces to determine stability. Test the hypothesis that learning can explain the variation observed over multiple trials. Propose non-linear objective function optimized by animal. Hypothesize feedback control laws.

Year 2: Vary sensory input from eyes and antennae and characterize the variability of individual insects to perturbations during running over multiple trials. Based on results, identify the sensory input used in learning. Implement non-linear objective function in two and three dimensional dynamic models. Test optimization hypothesis for stability. Propose updated feedback control laws.

Year 3: Test reinforcement learning model developed by team with dynamic models. Generalize learning model particularly with respect to the decentralized control apparent in insects.]]

C.9 Project management plan

[[SS update]]

The project brings together a very diverse group of researchers from integrative biology, psychology and electrical and computer engineering. Several of the PIs have worked together in group projects with each other in the recent past: for example, Russell and Sastry have worked on learning control for unmanned aerial vehicles on other projects, Full and Sastry ran some of the planning workshops for the LIS program together, Russell and Sastry are co-PIs on an ARO Multidisciplinary University Research Initiative (MURI) entitled "An integrated approach to intelligent systems". Full, Dickinson and Farley are co-investigators on several large locomotion projects. Full, Pister and Sastry have run seminars on control theory for locomotion in the past. Fearing, Sastry, Full, Dickinson and Pister are co-PIs on a ONR MURI on "Flexible Robots".

The PI and Director of the project with be Stuart Russell. He will obtain 25% release time (already approved) to manage the project. This includes a weekly seminar series, coordination of the test beds and coordination of the post doctoral scholars on the projects. The testbeds on flying winged insects, running cockroaches, running and throwing [[RI]] humans will provide the glue for the different approaches and the analytical models which will be proposed by team members. Part of the weekly seminar slots will be dedicated to status reports on the test beds, other slots will be for dedicated subject area presentations on learning theory, micro mechanical flying insects, locomotion, Bernstein's problems, aphasias, ataxias, stroke victim responses and so on. Post doctoral scholars on the project will be chosen to be co-supervised by at least two and preferably three of the team members to encourage synergy. We have good experience doing this on our large projects already. The project leader will highlight the annual accomplishments in the annual progress report which will also set goals for the next year as modifications of the original plans.

A web site with project publications, seminar series listings and announcements will be maintained for easy dissemination of project information. Full specializes in multi-media instructional tools for bringing the project research to undergraduate classrooms. We will leverage off this expertise of one of the team members to enhance the undergraduate research experience at Berkeley. It is our increasing conviction that research universities like us need to rope young scientists into the scientific research environment early and we will actively recruit talented undergraduates.

We have a record of training large numbers of some of the best graduate students and post doctoral scholars in the country from a diversity of ethnic backgrounds. The PIs have a strong record of mentoring minorities and women scientists. Our students have shown a preference to go to academia thus far, and we think that the students trained in this program will be academic material, but we have seen a surge in the number of attractive positions being offered to our alumni by the computer animation/graphics industry, as well as human machine interface companies.

D References cited

[[will become a separate section]]

References

Agarwal, M. (1997). A systematic classification of neural-network-based control. *IEEE Control Systems Magazine*, *17*(2), 75–93.

Alexander, R. M. (1988). *Elastic mechanisms in animal movement*. Cambridge University Press, Cambridge.

Alexander, R. M. (1992). A model of bipedal locomotion on compliant legs. *Phil. Trans. Roy. Soc.*, *B 338*, 189–198.

Allard, F., Graham, S., & Paarsalu, M. (1980). Perception in sport: Basketball. *Journal of Sport Psychology*, *2*, 14–21.

Athans, M., & Falb, P. L. (1966). *Optimal control: An introduction to the theory and its applications*. McGraw-Hill, New York.

Bellman, R. E. (1957). *Dynamic Programming*. Princeton University Press, Princeton, New Jersey.

Berns, G. S., & Sejnowski, T. J. (1996). How the basal ganglia make decisions. In Damasio, A. (Ed.), *Neurobiology of Decision-Making*, pp. 101–113. ??, ??

Bernstein, N. (1967). *The coordination and regulation of movements*. Pergamon, Oxford.

Berthier, N. E., Singh, S. P., Barto, A. G., & Houk, J. C. (1993). Distributed representation of limb motor programs in arrays of adjustable pattern generators. *Journal of Cognitive Neuroscience*, *5*(1), 56–78.

Bertsekas, D. C., & Tsitsiklis, J. N. (1996). *Neuro-dynamic programming*. Athena Scientific, Belmont, Mass.

Blickhan, R. (1989). The spring-mass model for running and hopping. *J. Biomechanics*, *22*, 1217–1227.

Blickhan, R., & Full, R. (1993). Similarity in multilegged locomotion: bouncing like a monopode. *J. Comp. Physiol.*, *A 173*, 509–517.

Bradtke, S. J., & Duff, M. O. (1995). Reinforcement learning methods for continuous-time markov decision problems. In Tesauro, G., Touretzky, D., & Leen, T. (Eds.), *Advances in Neural Information Processing Systems*, Vol. 8, pp. 393–400 Denver, CO. MIT Press.

Casey, T. M. (1981). A comparison of mechanical and energetic estimate of flight cost for hovering sphinx moths.. *J. Exp. Biol.*, *91*.

Chan, W. P., & Dickinson, M. H. (1996a). In vivo length oscillations of indirect flight muscles in the fruit fly Drosophila virilis. *J. Exp. Biol.*, *199*, 2767–2774.

Chan, W. P., & Dickinson, M. H. (1996b). Position-specific central projections of mechanosensory neurons on the haltere of the blow fly, Calliphora vicina. *J. Comp. Neurol.*, *369*, 405–418.

Chan, W. P., Prete, F., & Dickinson, M. H. (1998). Visual input to the efferent control system of a fly's 'gyroscope'. *Science*, *289*, 289–292.

Chase, W., & Simon, H. (1973). The mind's eye in chess. In Chase, W. (Ed.), *Visual Information Processing*, p. ?? Academic Press, New York.

Cohen, P., Morgan, J., & Pollack, M. (1990). *Intentions in Communication*. MIT Press.

Crawford, L. S., & Sastry, S. S. (1995). Biological motor approaches for a planar diver. In *Proceedings of the 34th IEEE Conference on Decision and Control*, pp. 3881–3886.

Crawford, L. S., & Sastry, S. S. (1996). Learning controllers for complex behavioral systems. ERL memo UCB/ERL M96/73, University of California at Berkeley. Department of Electrical Engineering and Computer Sciences.

Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, Oxford.

Dickinson, M. H. (1996). Unsteady mechanisms of force generation in aquatic and aerial locomotion. *Amer. Zool.*, *36*, 537–554.

Dickinson, M. H., & Götz, K. G. (1996). The wake dynamics and flight forces of the fruit fly, Drosophila melanogaster. *J. Exp. Biol.*, *199*, 2085–2104.

Dickinson, M. H., Hyatt, C. J., Lehmann, F.-O., et al. (1997). Phosphorylation-dependent power output of transgenic flies: an integrated study. *Biophys. J.*, *73*, 3122–3134.

Dickinson, M. H., & Lighton, J. R. B. (1995). Muscle efficiency and elastic storage in the flight motor of Drosophila. *Science*, *128*, 87–89.

Dickinson, M. H., & Tu, M. S. (1997). The function of Dipteran flight muscle.. *Comp. Biochem. Physiol. A*, *116A*, 223–238.

Dietterich, T. G. (1997). Machine-learning research: Four current directions. *AI Magazine*, *18*(4), 97–136.

Doya, K., & Sejnowski, T. (1995). A novel reinforcement model of birdsong vocalization learning. In Tesauro, G., Touretzky, D., & Leen, T. (Eds.), *Advances in Neural Information Processing Systems*, Vol. 8, pp. 101–8 Denver, CO. MIT Press.

Ellington, C. P. (1984). The aerodynamics of insect flight. VI. lift and power requirements. *Phil. Trans. R. Soc. Lond. B*, *305*, 145–181.

Ellington, C. P. (1985). Power and efficiency of insect flight muscle. *J. Exp. Biol*, *115*, 293–304.

Ellington, C. P. (1991). Limitations on animal flight performance. *J. Exp. Biol*, *160*, 71–91.

Ellington, C. P. (1995). Unsteady aerodynamics of insect flight. In Ellington, C. P., & Pedley, T. J. (Eds.), *Biological Fluid Dynamics*. Company of Biologists, London.

Ellington, C. P., v. d. Berg, C., Willmott, A. P., & Thomas, A. L. R. (1996). Leading-edge vortices in insect flight. *Nature*, *384*(6610), 626–30.

Ericsson, K., Krampe, R., & Tesch-Roemer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363–406.

Farley, C., & Gonzalez, O. (1996). Leg stiffness and stride frequency in human running. *J. Biomech.*, *29*, 181–186.

Farley, C., McMahon, T., & Glasheen, J. (1993). Running springs: speed and animal size. *J. exp. Biol.*, *185*, 71–86.

Fayyazuddin, A., & Dickinson, M. H. (1996). Haltere afferents provide direct, electrotonic input to a steering motor neuron of the blowfly, Calliphora. *J. Neurosci.*, *16*, 5225–5232.

Ferris, D., & Farley, C. (1997). Interaction of leg stiffness and surface stiffness during human hopping. *J. of Appl. Physiol.*, *82*(1), 15–22.

Ferris, D., Louie, M., & Farley, C. (1998). Running in the real world: Adjusting leg stiffness for different surfaces. Submitted to Proceedings of the Royal Society: Biological Sciences.

Franz, E., Eliassen, J., Ivry, R., & Gazzaniga, M. (1996). Dissociation of spatial and temporal coupling in bimanual movements of callosotomy patients. *Psychological Science*, *7*, 306–310.

Friedman, N., Murphy, K., & Russell, S. (1998). Learning the structure of dynamic probabilistic networks. In *Uncertainty in Artificial Intelligence: Proceedings of the Fourteenth Conference* Madison, Wisconsin. Morgan Kaufmann.

Full, R. (1993). Integration of individual leg dynamics with whole body movement in arthropod locomotion. In Beer, R., Ritzmann, R., & McKenna, T. (Eds.), *Biological Neural Networks in Invertebrate Neuroethtology and Robots*, pp. 3–20. Academic Press, Boston.

Full, R., & Ahn, A. (1995). Static forces and moments generated in the insect leg: comparison of a three-dimensional musculoskeletal computer model with experimental measurements. *J. Exp Biol.*, *198*, 1285–1298.

Full, R., & Farley, C. T. (1998). Musculoskeletal dynamics in rhythmic systems—a comparative approach to legged locomotion. In Winters, J. M., & E., C. P. (Eds.), *Biomechanics and Neural Control of Movement*. In press.

Gorinevsky, D., Kapitanovsky, A., & Goldenberg, A. (1996). Radial basis function network architecture for nonholonomic motion planning and control of free-flying manipulators. *IEEE Transactions on Robotics and Automation*, *12*(3).

Gottlieb, G. L., Corcos, D. M., & Agarwal, G. C. (1989). Strategies for the control of voluntary movements with one mechanical degree of freedom. *Behavioral and Brain Sciences*, *12*, 189–210.

He, J., McMahon, T. A., & Kram, R. (1991). Mechanics of running under simulated reduced gravity. *J. Appl. Physiol.*, *71*, 863–870.

Heglund, N., Fedak, M., Taylor, C., & Cavagna, G. (1982). Energetics and mechanics of terrestrial locomotion. IV. total mechanical energy changes as a function of speed and body size in birds and mammals. *J. Exp. Biology*, *97*, 57–66.

Helmuth, L., & Ivry, R. (1996). When two hands are better than one: Reduced timing variability during bimanual movements. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 278–293.

Ivry, R. (1997). Cerebellar timing systems. *International Review of Neurobiology*, *41*, 555–573.

Ivry, R., Franz, E., Kingstone, A., & Johnston, J. (1998). The prp effect following callosotomy: Uncoupling of lateralized response codes. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 463–480.

Kaelbling, L. P., Littman, M. L., & Moore, A. W. (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, *4*, 237–285.

Kahneman, D., Slovic, P., & Tversky, A. (Eds.). (1982). *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge.

Kawato, M. (1990). Feedback-error-learning neural network for supervised motor learning. In Eckmiller, R. (Ed.), *Advanced Neural Computers*, pp. 365–372. North-Holland, Amsterdam.

Kearns, M., Schapire, R., , & Sellie, L. (1992). Toward efficient agnostic learning. In *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory (COLT-92)* Pittsburgh, Pennsylvania. ACM Press.

Kearns, M. J., & Vazirani, U. V. (1994). *An Introduction to Computational Learning Theory*. MIT Press, Cambridge, Massachusetts.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Wiley, New York.

Knox, S., & de Chernatony, L. (1989). The application of multiattribute modelling techniques to the mineral water market. *Quarterly Review of Marketing*, *14*(4), 14–20.

Kram, R., & Taylor, C. (1990). Energetics of running: A new perspective.. *Nature*, *346*, 265–267.

L., M., & A., N. J. (1997). Oxygen consumption in the foraging honeybee depends on the reward rate at the food source. *Journal of Comparative Physiology. B, Biochemical, Systemic, and Environmental Physiology*, *167*(1), 36–42.

Lehmann, F.-O., & Dickinson, M. H. (1997). The changes in power requirements and muscle efficiency during elevated force production in the fruit fly, Drosophila melanogaster. *J. Exp. Biol*, *200*, 1133–1143.

Logan, G. (1982). On the ability to inhibit complex movements: A stop-signal study of typewriting. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 778–792.

Losos, J. B., & Mikels, D. B. (1995). Adaptation, constraint, and the comparative method: phylogenetic issues and methods. In Wainright, P. C., & Reilly, S. M. (Eds.), *Ecological Morphology*, pp. 60–98. University of Chicago Press, Chicago.

MacKay, D. G. (1987). *The Organization of Perception and Action.* Springer-Verlag, New York.

McGeer, T. (1990). Passive bipedal running. *Proc. R. Soc. Lond.*, *B 240*, 107–134.

McMahon, T. (1990). Spring-like properties of muscles and reflexses in running. In Winters, J., & y Woo, S. (Eds.), *Multiple Muscle Systems*, pp. 578–=590 New York. Springer.

McMahon, T., & Cheng, G. (1990). The mechanics of running: how does stiffness couple with speed?. *J. Biomechanics*, *23 (suppl. 1)*, 65–78.

Mochon, S., & McMahon., T. (1980). Ballistic walking. *J. Biomech.*, *13*, 49–57.

Mochon, S., & McMahon., T. (1981). Ballistic walking: an improved model. *Math. Biosci.*, *52*, 241–260.

Montague, P. R., Dayan, P., Person, C., & Sejnowski, T. J. (1995). Bee foraging in uncertain environments using predictive hebbian learning. *Nature*, *377*, 725–728.

Nalbach, G. (1993). The halteres of the blowfly Calliphora I. kinematics and dynamics.. *kinematics and dynamics. J. Comp. Physiol.*, *173*, 293–300.

Nevistic, V., & Morari, M. (1996). Constrained control of nonlinear systems: model predictive control based methods. *Automatica*, *37*(3-4), 105–9.

Papavassiliou, V., & Russell, S. (1998). Convergence of a new algorithm for reinforcement learning with general function approximators. submitted to NIPS-98.

Parr, R., & Russell, S. (1995). Approximating optimal policies for partially observable stochastic domains. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95)* Montreal, Canada. Morgan Kaufmann.

Pashler, H. (1998). *The Psychology of Attention*. MIT Press, Cambridge, MA.

Pil, A. C., & Asada, H. (1994). Recursive experimental structure re-design of a robot arm using rapid prototyping. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, Vol. 2, pp. 1094–1099.

Playter, R. R., & Raibert, M. H. (1994). Passively stable layout somersaults. In *Eighth Yale Workshop on Adaptive and Learning Systems*, pp. 66–71 New Haven, CT.

Pomerleau, D. A. (1993). *Neural Network Perception for Mobile Robot Guidance*. Kluwer, Dordrecht, The Netherlands.

Raibert, M., & Hodgins, J. (1993). Legged robots. In Beer, R., Ritzmann, R., & McKenna, T. (Eds.), *Biological neural networks in invertebrate neuroethology and robotics*, pp. 319–354 San Diego. Academic Press.

Russell, S. J., & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, New Jersey.

Sastry, S., & Bodson, M. (1989). *Adaptive Control: Stability, Convergence, and Robustness*. Prentice Hall, Englewood Cliffs, NJ.

Schmajuk, N. A., & Zanutto, B. S. (1997). Escape, avoidance, and imitation: a neural network approach. *Adaptive Behavior*, *6*(1), 63–129.

Strausfeld, N. J., & Lee, J.-K. (1991). Neuronal basis for parallel visual processing in the fly. *Visual Neuroscience*, *7*, 13–33.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. *Machine Learning*, *3*, 9–44.

Tesauro, G. (1992). Practical issues in temporal difference learning. *Machine Learning*, *8*(3–4), 257–277.

Thompson, C., & Raibert, M. (1989). Passive dynamic running. In Hayward, V., & Khatib, O. (Eds.), *International Symposium of Experimental Robotics*, pp. 74–83 New York. Spring-Verlag.

Ting, L., Blickhan, R., & Full, R. (1994). Dynamic and static stability in hexapedal runners. *J. Exp Biol.*, *197*, 251–269.

Tohtong, R., Yamashita, H., Graham, M., Haeberle, J., Simcox, A., & Maughan, D. (1995). Impairmaent of muscle function caused by mutations of phosphorylation sites in myosin regulatory light chain. *Nature*, *374*, 650–653.

Tomlin, C. J., Lygeros, J., & Sastry, S. (1998). Synthesis of safe controller for hybrid systems. In *Hybrid Systems: Computation and Control*. Springer Verlag.

Touretzky, D. S., & Saksida, L. M. (1997). Operant conditioning in Skinnerbots. *Adaptive Behavior*, *5*(3–4), 219–47.

Tsitsiklis, J. N., & Van Roy, B. (1996). An analysis of temporal-difference learning with function approximation. Technical report LIDS-P-2322, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology.

Tu, M. S., & Dickinson, M. H. (1996). The control of wing kinematics by two steering muscles of the blowfly, Calliphora vicina. *J. Comp. Physiol.*, *178*, 813–830.

Tu, M. S., & Dickinson, M. H. (1994). Modulation of negative work output from a steering muscle of the blowfly Calliphora vicina. *J. Exp. Biol*, *192*, 207–224.

Watkins, C. J. (1989). *Models of Delayed Reinforcement Learning*. Ph.D. thesis, Psychology Department, Cambridge University, Cambridge, United Kingdom.

Winter, D. A. (1995). Human movement: A system-level approach. In Arbib, M. A. (Ed.), *The Handbook of Brain Theory and Neural Networks*. MIT Press, Cambridge, MA.

Wright, C. (1990). Generalized motor programs: Reevaluating claims of effector independence.. In Jeannerod, M. (Ed.), *Attention and Performance XIII: Motor Representation and Control*, pp. 294–320. Lawrence Erlbaum, Hillsdale, NJ.