

Trends in Neuroscience, Vol. 6, #6, June, 1983

Finding common ground between robotics and physiology

SIR:

In recent years, it has been argued that motor control physiologists face a heuristic dilemma whose proportions are far beyond the theoretical problems of other areas of CNS function such as perception and memory. The usual argument begins by presenting two alternative solutions for controlling trajectory in a complex mechanical system: direct computation of the mechanical inverse and the use of look-up tables of pre-programmed movements⁴. The first of these straw men is then knocked down because the analysis leads to a Very Long Equation (VLE), which cannot possibly be computed by slow, noisy neural networks²; the latter is dismissed by the calculation of a Very Large Number (VLN) representing the size of the required memory, usually in excess of the number of protons in the universe.

At this point, the author usually offers a solution to the obvious crisis, which is the Simplifying Assumption (SA). SAs which have been offered over the past decade are generally constraints which fly in the face of our intuitive understanding of the capabilities and diverse strategies of real physiological motor systems, but which we are forced to entertain when we can offer no escape from the VLE and the VLN. Unfortunately, or perhaps fortunately, the promulgators of the SAs have found it relatively easy to disprove the universality of their creations. The length-servo⁵, the equilibrium spring-pair¹, and the Taylor series of derivatives⁷ have all been shown to have the sort of inadequacies in the real biological world which one might have guessed they would have. Interestingly, the only SA which has survived the test of time is the size principle for orderly motoneuron pool recruitment, which pre-dates the interest of roboticists in physiology.

To pursue the Aristotelian argument, one might conclude that we have not yet found the right SAs, but I prefer to attack the premises which got us to this point. If one is attempting to control the motion of a 12 axis robot to 0.001 cm anywhere in 100 cm of travel and that robot must leave the factory with the same or better specifications than it can be expected to have ten years later, then the discouraging analyses of the alternate general solutions seem incontrovertible. However, primate motor performance is nothing like that. The performance of an unrehearsed trajectory on command is, in

fact, rather poor. If practiced, it becomes rapidly more accurate³, but continues to have idiosyncracies which may be difficult to change. A specific motor task can be developed to extraordinary accuracy but the motor aspects of the skill do not generalize significantly to untrained parts. The motor controllers of the figure-skater and the violinist are specialized machines which have somehow become optimally tuned over time to the demands placed upon them.

All of these properties have much in common with the characteristics of perceptual systems such as vision and hearing. When confronted with a totally foreign scene, we do not draw a blank nor do we immediately perceive all its subtle nuances. Instead, we try some first approximations based on available similarities to previous experience. If we are forced to confront and interact with these new objects intensively, we soon learn to recognize subclasses based on progressively more subtle features. We can think of our perceptual space as a multidimensional set of axes with variably spaced bins whose widths correspond to our discriminative powers. Any discriminable entity occupies a unique place in this multidimensional space. The place is not point-like but has size depending on the grains of each axis in the region which projects to this part of space. The more frequently we encounter entities in one region of this multidimensional space, the finer we subdivide this part of space. Thus, the grain of our internal representation of our environment is influenced both by its physical reality and by our accumulated experience of the reality. However, the orientation of the dimensional axes tends to be fixed early in life, giving each individual his unique and idiosyncratic way of looking at things. While the perceptual resolution of two individuals equally practiced in an environment may become equal, the internal representation of the percepts is not identical. This use-dependent reshuffling of perceptual space may be the basis for the recently noted plasticity in cortical somatotopy⁶; it guarantees an even distribution of the work among neurons over time.

The performance of motor tasks can be interpreted within such a perceptual framework. When first confronted with a request to perform an entirely new movement, we look for similarities to previously performed movements and, perhaps, interpolate a reasonable approximation. If the movement lies within a previously well-explored domain, the similarities will be

close and the interpolated result quite good. If we have had little experience in this part of motor space, the movement will be inaccurate and we may fall back on strategies slow enough to facilitate conscious visual or proprioceptive feedback⁸. As we practice the movement, we quickly store newly tried motor programs, using a multidimensional representation based on apparent outcome and postural and other contextual information. This has the effect of crudely subdividing the space around the desired trajectory into smaller grain chunks to whatever limit we desire, constrained only by the resolution of our sense organs and motoneuron pools and willingness to practice. If this space then goes unused for prolonged periods, it is likely that the small chunks gradually coalesce back into larger ones permitting only less accurate performance. If we have indexed our motor space based on one set of axes (for example endpoints in extrapersonal space), then we may have difficulty learning to make accurate movements when the accuracy is judged by a criterion not mapped closely along any particular axis (for example velocity of rotation at a single joint). This hypothetical system makes no extraordinary demands on either memory or computational ability. Rather, it lives with its obvious limitations by constantly redistributing them in the least conspicuous manner. Instead of storing all possible programs or even all executed programs, it stores only the most useful programs. Parenthetically, this scheme is more compatible with a motor cortex organized by inputs (perceived trajectories required or experienced following previous output) rather than outputs (motoneuron pools), which casts new light on the old muscle vs. movement controversy.

Obviously, such a dynamic and indeterminate control system will be quite unattractive to the designers of robots, so it is not surprising that it has escaped their attention. However, its attraction to nature should be obvious. First, it directly supports the *raison d'être* of intelligent nervous systems, namely adaptability to situations which cannot be anticipated genetically. Second, it structures the motor performance problem into a form for which a very well developed computational mechanism already exists, namely perception. The major difficulty with this hypothesis is that it forces experimentalists to confront the very attributes of normal behavior which are most troublesome in an

experimental paradigm, namely the unreproducible, unpredictable meanderings of an adaptive control system working with an unrehearsed movement. Fortunately the study of motor skill acquisition has some methodological precedence in exercise physiology and rehabilitation medicine, but this has not been exploited by neurophysiologists thinking about the neural circuitry involved in motor performance or motor learning. Sensory physiology has profited immensely from asking questions about development, learning and conditioning; perhaps motor physiology can too. Ask not what your subjects can do, but how they learn to do it.

References

- 1 Bizzi, E., Chapple, W. and Hogan, N. (1982) *Trends NeuroSci.* 5, 395-398
- 2 Chandler, C., Hewitt, J. and Miller, S. (1982) *Trends NeuroSci.* 5, 376
- 3 Georgopoulos, A. P., Kalaska, J. F. and Massey, J. T. (1981) *J. Neurophysiol.* 46, 725-743
- 4 Hollerbach, J. M. (1982) *Trends NeuroSci.* 5, 189-192
- 5 Marsden, C. D., Merton, P. A. and Morton, H. B. (1972) *Nature (London)* 238, 140-143
- 6 Merzenich, M. M. and Kaas, J. H. (1982) *Trends NeuroSci.* 5, 434-436
- 7 Pellionisz, A. and Llinas, R. (1979) *Neuroscience* 4, 323-348
- 8 Woodworth, R. S. (1899) *Psychological Monographs*, Vol. 3, No. 13, pp. 1-114

GERALD E. LOEB

Laboratory of Neural Control, IRP, National Institute of Neurological and Communicative Disorders and Stroke, National Institutes of Health, Bethesda, MD 20205, USA.

ATTENTION

Neurology Residents in North America

Neurology residents in Canada and the US are eligible for the discount subscription rate advertised in this issue. Simply enclose evidence of registration with a recognised, university affiliated program for training in neurology and neurosurgery with your application, and send it to our New York Office.

Viewpoint

Learning - an evolutionary approach

Some behavioural variation in animals does not have a genetic basis. When such behaviour is flexible and adapts the animal to the environment, it is said to result from learning. Ethologists and psychologists have presented models of learning which are consistent with various behavioural patterns^{1,11,12}, but it has not been usual to ask whether these models could reasonably reflect the neurophysiological process which accounts for the learning, or whether such a process would be stable during evolution. If these questions are not answered, models of behaviour such as mathematical learning theory amount to little more than curve fitting. This report describes a model of learning which utilizes an approach based on both evolution theory and neurophysiology.

Choice behaviour: learning a game

We assume that in animals capable of learning, a neural process called a 'learning rule' produces responses which vary, usually in an adaptive manner, with the rewards or 'payoffs' associated with the responses. However, learning can range from classical conditioning to creative or insight learning⁹ and it is unlikely that the underlying neural processes are identical across this spectrum. We will consider only the learning of a 'game', i.e. a series of situations where (i) a choice of mutually exclusive behaviours is available at each trial, and (ii) a payoff is assigned to each situation in a defined manner.

To simplify the analysis, we have made several assumptions which to us seem justified. Consider a two-choice game. We assume that the learning rule defines the probability of displaying each behaviour at each trial. These probabilities sum to 1. We also assume that animals associate behaviours with the contingent payoff, and that they can evaluate payoffs of different quantity or quality in a common unit related to biological fitness. The capacity for such an evaluation probably evolved prior to the capacity to learn, since the ability to assess the value of many primitive actions according to physiological need¹⁰ seems ubiquitous. However, it is important to note that the ability to evaluate payoffs and associate them with the appropriate behaviour cannot be perfect. In some situations, especially artificial laboratory games, an individual could easily misinterpret payoffs.

Our first attempt to deduce the nature of the learning rule for choice behaviour is based on the concept of evolutionary stability.

Evolutionary stability

Populations of animals evolve by natural selection acting on heritable variations in phenotype which affect survival or reproductive potential. Biologists measure the

relative survival of the determinant of the phenotype (usually a gene or combination of genes) in units of 'fitness'. In addition to intrinsic properties of the phenotype, fitness will depend upon the population and environment.

Fig. 1 illustrates the evolution of a population composed of two phenotypes, 'hawk' and 'dove', whose fitness varies in a particular frequency-dependent manner. This dependence is modelled after the hawk-dove game¹³⁻¹⁵ in which individuals compete for a resource, and the probability of winning the resource depends upon the strategy of both contestants.

It is apparent that after some generations the population stabilizes at a frequency of hawk and dove where the fitness of these two phenotypes is identical. It is important to note that even though the overall fitness of the population is greatest when the frequency of hawks is 0, the population is not 'evolutionarily stable' at that point: hawks can invade a predominantly dove population because of their superior fitness under those conditions. Analysis of games such as this lead to a precise definition of evolutionary stability.

Evolutionary stability refers to a phenotype which has the following property: when members of a population possess the evolutionarily stable (ES) phenotype, a mutant with any alternative phenotype within an allowable range is less fit than those with the ES phenotype¹⁶⁻¹⁷. In other words, the ES phenotype is evolutionarily stable since whenever mutant phenotypes arise they cannot successfully invade the population. A polymorphism, i.e. a mixture of phenotypes within a population, may also be evolutionarily stable under certain circumstances. In the example above, neither pure hawk nor pure dove is evolutionarily stable, but both a mixed strategy ($3/4$ hawk, $1/4$ dove) and a polymorphism are stable.

Evolutionary stability was originally