

currency for comparing perhaps qualitatively different activities; it also uses an explicit dynamic programming framework defined over actions within a specified time interval. I briefly discuss these two key features applied to agents who are *imperfect* at always selecting the dynamically optimal sequence of actions over a given time interval. Such imperfect agents will thereby repeatedly select actions with *positive* canonical costs. Perfect agents, on the other hand, would always select actions with zero canonical costs. Thus, H & M's methodology seems naturally connected to imperfect decisions that allow positive canonical costs to arise.

As currently defined, the canonical cost of a particular action at a given point in time is the loss compared to selecting the best action at that time, where the best action is defined assuming all subsequent actions over the time interval are optimally chosen. If, however, agents may select nonoptimal actions at some given point in time (because of finite capacity decision skills, thinking and perceptual apparatus, etc.), then some of their subsequent actions may also be chosen imperfectly. Consequently, it may be inappropriate to define an opportunity cost relative to an action whose optimality depends on the behaviorally infeasible possibility of choosing all subsequent actions perfectly. In my related work on imperfect choice theory (see H & M's citations and Heiner 1985a), it has been helpful also to define loss measures for decision errors relative to an agent's *attainable* expected future performance, conditional on having the same decision and perceptual abilities for future decisions present when making any particular current decision. Defining loss measures for decision errors as described also enables one to investigate the comparative behavioral effects (say, across different species) of changing an agent's cognitive and perceptual mechanisms instead of only defining canonical costs as if a hypothetically perfect agent were making future decisions.

The above comments are not critical of the basic methodology of canonical costs, but rather suggest that broadening their definition may permit a deeper analysis of agents with imperfect decision skills.

My second point concerns the "backward induction" dynamic programming framework used to define the optimal sequence of actions over a given time interval. As H & M stress, optimal decisions may depend on an animal's state at any point in time, which may itself vary over time in a complex and stochastic manner. As they further note, best current actions may further depend on the sequence of future actions. It is also possible to incorporate additional nonlinear dynamic links that can make each next optimal decision depend on the entire past history of decision states, as well as producing mathematically "chaotic" optimal sequences. These possibilities amplify the general point that once properly modeled, optimal dynamic behavior may depend on extremely complex determinants and may itself be extremely complex over time.

This recognition complements a point already noted by H & M, namely, that, realistically speaking, animals cannot solve these complex dynamic optimality relations, and thus must somehow resort to much simpler decision rules. My suggestion here is that H & M should go beyond this intuitive conclusion to analyze formally how imperfect ability to behave in the complex manner required for dynamic optimality can be used to derive more precisely the form that "robust" decision rules may take, and what sort of (perhaps instinctive) behavioral mechanisms will be evolutionarily viable.

H & M have already indicated one such possibility: High reliability may be required for an animal to benefit from adopting risk-prone behavior. Consequently, selection pressure may be created in favor of "risk-averse" strategies intended to be robust only over frequently (rather than rarely) encountered circumstances.

A specific instance of this mentioned by H & M is the case of male sticklebacks "attacking objects with red underparts" (since red underparts is a crude but still correct signal for attacking

under most naturally encountered situation(s)). Many other cases could be explored, including the general possibility of dynamically stable behavior itself arising from the need to control decision errors over time (e.g., Heiner, in press a). It may thus be fruitful to combine the methodology of canonical costs, defined using dynamic programming concepts, with explicit analysis of the behavioral effects that are induced in order for imperfect agents to control decision errors. These decision errors concern how to "adapt" to different circumstances and what to "learn" from past experience.

Optimization theory: A too narrow path

Gene M. Heyman

Medical Research Division, Lederle Laboratories, Pearl River, N.Y. 10965

Evolution is a historical process. The effects of environmental selection are conditioned by the available genetic variation, and the range of genetic variation depends in part on previous environmental challenges. Moreover, the stamp of history has not been faint. The vast majority of species that have appeared on earth have died out, and this occurred because genotypes simply did not change at the rate demanded by environmental pressures. An implication of adaptive constraints is that pervasive traits need not be optimal. Indeed, all that is necessary, according to the logic of evolution, is a relatively successful adaptive outcome. However, Houston & McNamara (H & M) assume that inherited behavior is necessarily optimal. Animals forage, they claim, in the way that maximizes reproductive success. Yet there are well-established instances in which animals did not gather foodstuffs in an economically rational way (Houston 1986). H & M acknowledge such discrepancies, but attribute them to limits in the animals rather than to limits in their theory. For example, if an animal did not maximize energy gain, H & M suggest the hypothesis that its behavior was constrained by inherited, simple "robust decision rules" that are optimal in some but not all situations. Of course, it is possible that in some instances environmental pressure and genetic variability are sufficient to yield an animal that forages optimally – well enough, in effect, to overcome its history – but this chain of events is not inevitable. Consequently, the logic of evolution requires us to entertain the possibility of both optimal and nonoptimal adaptations. In line with this logic, observations from operant psychology and behavioral ecology show pervasive, stable nonoptimal behavioral patterns.

Researchers in operant psychology have been concerned with how animals allocate behavior (choice). In one widely used procedure, the subjects are provided with two or more simultaneously available reinforcement sources. The reinforcements are delivered intermittently and subjects respond at both sources, switching back and forth. The response totals at the end of the session are quite reliable: Relative response rates (choice) approximate relative reinforcement rates (e.g., relative number of items consumed per hour). This relationship is known as the matching law (Herrnstein 1970), and because of its generality it has attracted the interest of researchers in economics (Heiner 1983) and behavioral ecology (Houston 1986). However, matching is not the pattern of choices that maximizes reinforcement rate. H & M acknowledge this, and they try to show that although matching does not maximize food intake it does maximize survival (which appears to be synonymous with reproductive success). To evaluate their analysis, we need to review some of the basic facts concerning matching.

Herrnstein introduced the matching principle in a paper (1961) that described the results of a study in which pigeons had access to two simultaneously available reinforcement sources. Each reinforcement source included a button, which when pecked occasionally operated a feed hopper. Herrnstein varied

the likelihood of feeding opportunities and found that a simple equation summarized the relationship between relative response and relative reinforcement rates:

$$\frac{B_1}{B_1 + B_2} = \frac{R_1}{R_1 + R_2}$$

where B_i is the frequency of responding at alternative i and R_i is the frequency of reinforcement at alternative i . It is important to note that in this procedure wide fluctuations in the distribution of choice proportions produced small changes in the reinforcement proportions, so that the range of possible quantitative relations was large. In other words, matching was not due to constraints imposed by the apparatus.

Following Herrnstein's report, researchers sought to establish the domain and functional significance of the matching relation. Different species were tested, including humans, rats, and pigeons, and the results revealed no signs of species-dependent constraints. Different reinforcers were tested, and the matching law applied equally well to different motivational systems (the reinforcers have included food, water, money, social approval, and brain stimulation). Different response topographies were tested, and the matching law described discrete responses, such as a lever press, and continuous periods of activity, for example time at a task, equally well (see de Villiers, 1977, for a review).

There has also been interest in whether matching applies outside the laboratory to natural settings. Baum (1974) and Houston (1986) showed that it does. In Baum's study the subjects were a flock of pigeons that roosted in the attic of an urban home. In Houston's study the subjects were a number of pied wagtails, insectivorous birds, that divided their foraging activities between a river bank and an open field. In both studies the subjects were not constrained, and Baum and Houston pointed out that a wide range of possible quantitative outcomes was possible. Yet, as in the laboratory, the distribution of behavior approximated the distribution of reinforcements.

The empirical generality of the matching relation and the assumption that behavior must in some sense be optimal led to a series of studies aimed at discovering just what was maximized when animals matched (Shimp 1969; Rachlin et al. 1976). [See also Rachlin et al.: "Maximization Theory in Behavioral Psychology" *BBS* 4(3) 1981.] In some studies economies were arranged in which matching and maximizing (either food or net energy gain) predicted different outcomes (DeCarlo 1985; Herrnstein & Heyman 1979; Heyman & Herrnstein 1986). Without exception, matching provided the more accurate account of the animals' behavior. That is, the subjects did not maximize food or net energy gain. Animals in a natural setting showed the same result. The pied wagtails could have caught more spiders and insects had they deviated from matching, but matching nevertheless persisted (Houston 1986).

H & M acknowledge that matching is incompatible with energy maximization, but they say that this is not necessarily at variance with their position. They write, matching makes "evolutionary sense for animals that forage in groups," by which they appear to mean that the costs and benefits of foraging combine with those of defending against competitors so as to yield matching. This is an important point, because if H & M are unable to account for matching, they are unable to account for one of the most general phenomena in the study of reinforced behavior.

However, it is not clear why H & M think that competition in foraging explains matching, because they offer no evidence to support their claim. I am not aware of studies on competition and matching, but the following results seem relevant to the issue. The matching law describes behavioral allocation equally well in pigeons, rats, and humans. These species do not engage in foraging to the same degree, and when they do forage, they do so for different items, under different metabolic constraints, and under different degrees of competition. Thus it seems unlikely

that competition between foragers will explain matching or why animals fail to maximize food intake. That different species match also argues against the evolution of matching as a simple, robust decision rule that approximates the optimal solution, for it is unlikely that the same "short-cut" would have independently evolved in different species. Houston (1980) expressed a similar conclusion when he wrote that the evolution of approximately optimal, simple decision rules implies that "Different species can be expected to differ in the problems they can solve" (p. 299).

H & M point out that preference for different foods and different activities will vary as a function of deprivation, time of day, season, and other factors. This is an important issue and one that has apparently received little attention in previous studies of foraging. However, it is not clear that H & M's approach will account for state- and time-dependent changes in foraging. For example, Snyderman (1983) showed that at lower body weights, pigeons were more likely to choose a small immediate reward than a large delayed one. This strategy decreased overall rate of food intake, so that on the face of it the pigeons' choices did not seem to maximize their likelihood of producing offspring. However, H & M point out that the pigeons' increased preference for an immediate reward may have increased their likelihood of surviving a short period of time, which, in turn, could positively affect reproductive success. However, data on deprivation, foraging strategy, and number of offspring were not presented, so that it is not possible to tell whether H & M's account is correct.

That evolution is an historical process means that stable and widespread behaviors can be nonoptimal. However, Houston and McNamara's approach does not make sense of nonoptimal adaptive behavior. It thus seems too narrow a framework for understanding foraging behavior.

The functional analysis of behaviour: Making room for Prufrock

Felicity A. Huntingford and Neil B. Metcalfe

Zoology Department, Glasgow University, Glasgow G12 8QQ, Scotland

The vision that gave rise to the first optimality models also heralded the birth of behavioural ecology. In the interests of tractability, the early models were simple; since then they have been made more complex and realistic, in the light of empirical and theoretical studies. The target paper by Houston & McNamara (H & M) continues this trend. In particular, it is concerned with the ways in which long sequences of behavioural acts, rather than single decisions, can be optimised. The placing of decisions in a time-dependent context is a clear step forward from the more conventional approach that treats each decision in isolation.

A second development is H & M's use of fitness (here equated with expected future reproductive success) as the scale against which alternative strategies are weighed. Classical optimality models, although acknowledging that fitness is the ultimate parameter upon which selection acts, have settled for the optimisation of more proximate parameters such as the net rate of energy gain. In contrast, the approach of H & M considers the animal's state at time T to map directly to fitness by means of a reward function. Thus, in theory, the fitness consequences of all behavioural acts can be calculated through their effect on the animal's state at time T .

From this it follows that the approach also allows a comparison of the merits of different options at each decision-point, through an examination of their effect on future fitness. This has two important benefits. First, we can calculate the strength of the selection pressure against suboptimal choices and so identify those decisions that have a major impact on fitness. Second, it