

Protean Behavior in Dynamic Games: Arguments for the co-evolution of pursuit-evasion tactics

Geoffrey F. Miller and Dave Cliff

School of Cognitive and Computing Sciences
University of Sussex, BRIGHTON BN1 9QH, U.K.

Email: geoffm, davec@cogs.susx.ac.uk

Abstract

Animals often chase each other about. These pursuit-evasion contests require the continuous dynamical control of complex sensory-motor behavior, and give rise to some of the most common and challenging co-evolutionary arms races in nature. This paper argues for the importance and fruitfulness of studying pursuit-evasion scenarios using evolutionary simulation methods. We first review the biological ubiquity of pursuit-evasion contests and the *protean* (adaptively unpredictable) behavior that often evolves in evasion strategies (e.g. when prey zig-zag to evade predators). We then review the *differential game theory* relevant to analyzing pursuit-evasion games, including the proven optimality of *mixed strategies* (corresponding to protean behavior) in many such games. Previous simulation work on evolving pursuit-evasion tactics is also reviewed. Following this, we describe results from some initial experiments that extend previous work in evolutionary robotics to explore the co-evolution of pursuit and evasion tactics in populations of simulated robots. We conclude with some possible scientific implications and engineering applications.

1 Introduction

Contests of pursuit and evasion are among the most common, challenging, and important adaptive problems that confront mobile animals, and are some of the most important potential applications for robots and other artificial autonomous agents. In a typical contest of this sort, a predator chases a prey animal around until the prey is eaten or the predator gives up. More symmetrically, two members of the same species may fight over a territory or resource, alternating between attack and defense tactics analogous to pursuit and evasion tactics. Pursuit and evasion behaviors, like attack and defense behaviors, tend to co-evolve against one another, resulting in some of the most intense and sustained evolutionary arms races in nature. Although pursuit-evasion (P-E) contests have been relatively neglected in research on the simulation of adaptive behavior, they have five major features that render them interesting and relevant.

First, pursuit and evasion strategies require highly robust forms of adaptive behavior and have particularly important fitness consequences. Animals that pursue or evade must maintain complex sensory-motor coordination with respect to both a physical environment and a hostile animate opponent. Pursuit-evasion contests also require continuous, real-time, dynamical control, in the face of an opponent that will ruthlessly exploit any delay, uncertainty, or error. Natural or artificial behavior-control systems that are slow, brittle, easily confused, or error-prone do not survive long in P-E scenarios. For these reasons, traditional artificial intelligence methods may prove particularly poor as models of P-E behaviors, and newer reactive, behavior-based, bottom-up approaches (e.g. [32]) may prove particularly apt.

Second, pursuit and evasion strategies evolve against one another in an ongoing, open-ended, frequency-dependent way, so P-E contests often give rise to co-evolution within or between species. Because P-E scenarios may be the simplest and most common cases of behavioral co-evolution, their investigation may illuminate behavioral arms races in general (see [23]). Such sustained co-evolution reinforces all of the challenges discussed in the previous paragraph: temporary adaptive advantage is continually eroded under co-evolution as new tactics arise. Co-evolution probably drives the evolution of both special perceptual capacities to entrain, track, and predict animate motion, and special motor capacities to generate complex, robust, unpredictable behavior [40]. Understanding both motion perception and motor control may thus depend on appreciating the role of P-E contests in behavioral evolution.

Third, P-E contests have received serious attention from at least three scientific disciplines: behavioral biology, neuroethology, and game theory. Animal behavior studies have revealed the ubiquity and importance of P-E tactics, anti-predator behaviors, and fighting skills [14]. The centrality of such behaviors is revealed by the fact that P-E games are the most common form of animal play behavior; such play facilitates learning sensory-motor coordination through "developmental arms races" between play-mates. Neuroethology (e.g. [6]) has spent much effort understanding neural systems for pursuit ("approach") and evasion ("avoidance"), including: ex-

plorations of specific circuits for rapid startle and escape behaviors (e.g. [7]); the role of the (very fast) tectospinal pathway in mediating pursuit behaviors in vertebrate predators such as salamanders, frogs, cats, and owls (e.g. [2]); and the specific attunement of sensory cells to patterns of animate motion relevant to pursuit and evasion (e.g. [3]). Game theorists have also studied P-E contests intensely for several decades, because of their importance in tactical air combat (e.g. telling pilots how to evade guided missiles) and other military applications (see [55]). "Differential game theory" [27] has developed a vocabulary for analyzing the structure and complexity of P-E games, and a number of formal results concerning optimal strategies for particular P-E games. We review the relevant animal behavior studies and game theory at length in the next sections.

Fourth, P-E contests are of considerable theoretical interest because they give rise to several unique behavioral phenomena. They are the simplest situations that can favor "protean" (adaptively unpredictable) behavior, as when prey animals zig-zag unpredictably to escape predators (see e.g. [8, 14]). Further, because effective pursuit may often require prediction and 'mind-reading', while effective evasion may require the use of unpredictable or deceptive tactics [14], such contests raise issues of signaling, communication, and tactical deception [38], and may provide a natural bridge from the evolution of basic sensory-motor control to the evolution of social psychology and 'Machiavellian intelligence' [5, 38].

Fifth, the study of P-E behaviors has many scientific implications and practical applications. A better understanding of the evolutionary, behavioral, and cognitive dynamics of P-E contests would have wide implications for animal behavior, neuroethology, comparative psychology, and evolutionary psychology [40]. Better methods for evolving P-E behaviors would have many applications in robotics, video games, virtual environments, and any other technology where real or simulated mobile agents come into behavioral conflict with other agents.

Because P-E contests are a major new area of investigation for simulation of adaptive behavior, which we hope will excite much further research, this paper lays out the biological and game-theoretic foundations in some detail: this is essentially a position paper. We are currently conducting experiments that extend previous work in evolutionary robotics [10] to investigate: (1) whether co-evolution between simulated robots engaged in P-E contests can lead to the more and more complex pursuit and evasion tactics over generations; (2) whether the use of continuous recurrent neural networks as control systems allows the emergence of more interesting and dynamic perceptual, predictive, pursuit and evasion abilities; and (3) whether the incorporation of random-activation units in the control system allows the evolution of adaptively unpredictable tactics. Section 5 discusses some prelim-

inary results. For a longer version of this paper, with a more extensive bibliography, see [39]. For a more detailed discussion of our methods and results, see [11].

2 Biological foundations

2.1 *The generality of pursuit-evasion problems*

We use the term "fitness affordances" (FAs) to denote things that have particular statistically expected consequences for the fitness (survival or reproduction) of organisms of a particular age, sex, condition, and species [37, 40, 50]. Positive FAs such as food, mates, shelter, or offspring have a positive expected effect on the replication of one's genes; negative FAs such as poisons, predators, parasites, and physical dangers have a negative expected effect. FAs are objective features of the environment insofar as their potential fitness effects exist regardless of whether the organism facing them knows or cares of their existence, but FAs are relational insofar as their biological significance exists only in relation to organisms with particular modes of survival and reproduction (e.g. what is food to one species is poison to another in a perfectly objective and yet perfectly relational way). Most FAs are spatially localized (at some scale), and only impose their fitness effects on organisms immediately present.

Mobility allows animals to actively approach and exploit positive FAs, and to actively avoid negative FAs. But very often, FAs are themselves mobile animals with their own fitness interests, which may turn the problem of approach into a problem of pursuit, or the problem of avoidance into one of evasion. Whenever there are recognized conflicts of interest over fitness effects between mobile animals, we may predict manifest conflicts of movement. If the conflict of movement is sustained across some expanse of space and time, there is a P-E conflict. Thus, interactions between two mobile agents that have conflicting expected fitness effects creates a pursuit problem for one agent and an evasion problem for the other.

The above scenario, where the roles of pursuer and evader are pre-determined by the FA relationships, and fixed for the duration of the interaction, could be called an asymmetric P-E contest. More symmetric sorts of P-E contests can unfold when similar animals both seek access to a positive FA, such as a territory or resource, that is worth more if enjoyed alone. Animals of the same species often fight over food, nest sites, and mates. In symmetric contests, the roles of pursuer and evader can switch back and forth rapidly as the animals take offensive or defensive roles. In both symmetric and asymmetric contests, the immediate behavioral conflict can result in a co-evolutionary arms race between pursuit and evasion tactics, if the agent-types in question encounter each another with reasonable frequency and with significantly

opposed fitness consequences across generations.

2.2 Biological observations

Pursuit is fairly simple: animals are usually observed to move towards the remembered, observed, or predicted location of the target. Evasion is more complex. For example, animal escape behavior in asymmetric P-E contests generally breaks down into three phases: (1) directional fleeing if a predator (or other negative, mobile FA) is threatening but still distant; (2) erratic zig-zagging if the predator begins catching up; and (3) convulsive 'death-throes' if caught. Directional fleeing is about as simple as directional chasing, but the last two tactics, zig-zagging and convulsing, are examples of a more interesting type: protean behavior.

Animals generally evolve perceptual and cognitive capacities to entrain, track, and predict the movements of other biologically-relevant animals such as prey, predators, and potential mates [40]. Such predictive abilities mean that unpredictable behavior will inevitably be favored in many natural P-E situations. For example, if a rabbit fleeing from a fox always chose the single apparently shortest escape route, the very consistency of its behavior would make its escape route more predictable to the fox, its body more likely to be eaten, its genes less likely to replicate, and its fitness lower. Predictability is punished by hostile animals capable of prediction. This is the basic logic behind the theory of protean behavior: the effectiveness of almost any behavioral tactic can be enhanced by endowing it with characteristics that cannot be predicted by an evolutionary opponent [14]. An arms race between perceptual capacities for predicting animate motion, and motor capacities for generating protean behavior, will generally result from evolutionarily recurring P-E contests [40].

Along with directional fleeing, protean escape behaviors are probably the most widespread and successful of all behavioral anti-predator tactics, being used by virtually all mobile animals on land, under water, and in the air. Driver and Humphries [14] review examples from hundreds of species, including humans. Predators can also exploit unpredictability to confuse prey, as when weasels do "crazy dances" to baffle the voles they stalk.

Even if erratic zig-zagging fails, another form of proteanism, convulsive behavior, may succeed. Sudden, unpredictable, vigorous "death-throes", alternating with puzzling passivity ("playing dead") is often effective at allowing prey to escape from predators [14]. Adaptive convulsions can also occur in more abstract state-spaces, as when cuttlefish and octopi undergo rapid color changes to defeat the search images (perceptual expectations) of their predators. Additional confusion effects may arise from group flocking and mobbing behaviors that include unpredictable movements, complex motion dynamics, and confusing coloration (zebra stripes or

shiny scales on fish). Unpredictability can also be exploited by divergence between individuals, as when animals within a species evolve "aspect diversity" (polymorphic coloration or behavior) through "apostatic selection" that favors low-frequency traits (e.g. because predators' use of search images penalizes common appearances). Indeed, apostatic selection may be a general feature of P-E arms races: novel and unexpected tactics may be favored at a variety of levels.

Co-evolution itself can be viewed as a P-E contest, operating between lineages rather than between individuals. From this perspective, sexual recombination makes sense as a protean strategy which unpredictably mixes up genes so as to "confuse" pathogens [22]. Indeed, this proteanism argument is one of the leading explanations for the evolution of sex [45]. Despite proteanism's importance, it has been long overlooked in biology, because complex order rather than useful chaos was assumed to be the defining feature of Darwinian adaptations.

3 Game-theoretic foundations

3.1 Differential pursuit-evasion game theory

Game theory [44] is concerned with the formal analysis of situations called "games" where: (1) players can choose different strategies that determine their actions under particular conditions; (2) conditions and outcomes unfold through the interactions of the players' strategies; and (3) players have preferences among outcomes, i.e. payoffs exist. In brief, players are agents that can make choices, implement strategies, and receive payoffs.

Traditional game theory focused on games with discrete moves (e.g. chess), but in the 1950s, Isaacs (e.g. [27]) wondered whether game theory could be used to model P-E situations such as aerial combat, where moves unfold continuously over time. Isaacs had two basic insights. First, P-E contests do require game theory rather than simple optimality theory, because the optimal pursuit strategy for one player (e.g. a guided missile) depends on the evasion strategy adopted by the other player (e.g. an aircraft), and vice-versa. This chicken-and-egg problem is precisely what game theory is good at analyzing. Second, the continuous nature of P-E contests can be modeled using differential equations that specify how state conditions (such as player positions and velocities) change incrementally as a function of players' strategies and previous state conditions: pursuit and evasion moves become continuous trajectories through a state-space. Isaacs [27] developed the "Tenet of Transition" which specifies that players must optimize (find the minimax solution for) the transitions between states leading towards a goal-state, which can be represented as optimizing the temporal derivatives of the relevant state variables. For example, pursuers try to minimize the time until capture and evaders try to maximize it. Ap-

plying the tenet of transition, pursuers at each moment in time should try to maximize the rate of their instantaneous approach towards the capture-state, and evaders should try to minimize it. If a solvable set of differential equations can be written that specify the continuous effect of strategies on state-conditions, then the optimal pursuit and evasion strategies can be found by applying the tenet of transition. Isaacs's ideas have proven enormously fruitful: Rodin's [46] recent bibliography of P-E differential game theory contains about 1200 entries, and theoretical results are often used in practical situations such as design of military aircraft control systems.

Differential P-E games are defined by a set of controls (what each player can do), a set of dynamics (that maps from the control variables onto the state variables of the game, and from state variables at one moment in time to the next moment), and a set of termination conditions (state conditions that determine when successful capture or evasion happens). For example, in a classic case analyzed by Berkovitz [4], a pursuer and an evader move with equal and constant speed in a plane, and control the direction of their velocity vector (which thus becomes their control variable). These two velocity vectors give rise to a system of first-order differential state equations that determine how the players move over time. The pursuer wants to minimize time to capture the evader and the evader wants to maximize time until capture, with capture defined as proximity within some small distance. Both players know the present state of the game (e.g. both of their positions and velocity vectors) but at each time-point they make separate and simultaneous decisions about what to do next. The available strategies are therefore functions that map from current states of the game (i.e. the positions and velocity vectors of both players) onto velocity-vector decisions about what direction to move next. In all differential games, strategies determine trajectories through the relevant state-space; in P-E games, strategies determine trajectories through physical space. From each player's perspective, the game becomes a problem of optimal spatio-temporal control with respect to the opponent and the environment. Indeed, control theory can be viewed largely as the solution of one-player differential games [28]; differential game theory addresses the more complex multi-player cases.

In classic "asymmetric" games (e.g. missile vs. aircraft), the roles of pursuer and evader are pre-determined and fixed. But in "symmetric" games (e.g. aircraft vs. aircraft), both players can collect payoffs for successful pursuit and successful evasion. Symmetric P-E contests have been analyzed as "two-target games" [18, 36]. The symmetric contests in our initial experiments (see [11]) resembled the sort of two-target aerial combat games that have been subject to intense game-theoretical analysis for several decades (see [21] for review).

3.2 The optimality of mixed strategies

The key to formal analysis in game theory is for games to be reduced from descriptive form (e.g. rules and heuristics) or "extensive form" (i.e. decision-tree form) to "normal form" (i.e. a joint payoff matrix that lists game outcomes given all possible strategies for all players). Some games in normal form have "minimax solutions" (a.k.a. "saddle points") that minimize each player's expected loss regardless of what the opponent does to maximize their expected gain; minimax solutions, if they exist, are jointly optimal for rational players. In games of perfect information, players are precisely and continuously aware of all moves made by other players, so that deception, confusion, and uncertainty are impossible. All games of perfect information have one or more saddle points corresponding to "pure" deterministic optimal strategies (though finding them may often be difficult, as in chess).

However, games of imperfect information (e.g. games where deception is possible) may have multiple saddle points or no saddle points. In such cases, "mixed strategies" (probability distributions across pure strategies) may be optimal. Perhaps the most important result from [44] was that every two-player, zero-sum game of incomplete information with multiple saddle points has an optimal strategy that is mixed rather than pure:

"One important consideration for a player in such a game is to protect himself against having his intentions found out by his opponent. Playing several different strategies at random, so that only their probabilities are determined, is an effective way to achieve a degree of such protection. By this device the opponent cannot possibly find out what the player's strategy is going to be, since the player does not know it himself. Ignorance is obviously a good safeguard against disclosing information directly or indirectly." [44, p.146]

The logic of mixed strategies is simple. If a player's choice sometimes remains unknown to others after the move is made, the game is one of imperfect information. This can result from the move being hidden, or the other players' sensors being insufficient to register all moves with complete accuracy. Typically, games lose their saddle points when they are no longer games of perfect information, such that the first player's minimax solution does not correspond to the second player's minimax solution. For example, the popular children's game Rock, Paper, Scissors involves a circular pattern of dominance among the pure strategies (Rock beats Scissors, Scissors cut Paper, Paper smotheres Rock), so there is no saddle point, and one's optimal (minimax) strategy against a rational opponent is to choose each move with one-third probability. In general, mixed strategies randomize moves to confuse opponents and keep them guessing.

(But the task of determining the optimal mixed strategy is usually very difficult for games with many pure strategies and complex interactions.)

Because many P-E games are games of incomplete information with multiple saddle points, mixed strategies have often proven useful in such games. Mixed strategies are optimal for a P-E game with rectilinear movement on a planar grid [15]. In some more complex continuous cases, the optimal strategies for both pursuer and evader are also mixed. Important recent work in this area has been by Forte and Shinar (e.g. [16, 49]); they showed that in aerial combat scenarios, mixed strategies yielded much better performance than any previously known guidance law, and did so for both pursuers and evaders. Such game-theoretic results support the protean behavior hypothesis of Driver and Humphries [14] that erratic zig-zagging by animals is truly stochastic behavior that derives its utility from its unpredictability. We might expect then that in any P-E game with incomplete information and complex dynamics, unpredictable pursuit and evasion strategies will evolve.

Evolutionary game theory [33] has also recognized the optimality of mixed strategies in many contests between animals. Animals can be considered players in the game-theoretic sense because they make choices, implement behavioral strategies, and receive fitness payoffs contingent on their interactions with other animals' strategies. Mixed strategies can be implemented as behavioral polymorphisms across individuals in a population or as protean behavior within each individual. However, evolutionary game theory has focused mostly on single-step games (such as sex-ratio determination or the Hawk-Dove game: see [33]) and discrete-step games (such as the iterated prisoner's dilemma). The literature on differential P-E games has been strangely overlooked despite its obvious relevance to predator-prey interactions and territorial fights, so the importance of protean evasion behavior has been neglected. Dynamic programming methods (e.g. [25, 26]) may prove more useful in analyzing P-E contests, since they can optimize stochastic dynamic strategies, even in two-player games (e.g. [9]). However, such methods require the specification of a fairly well-defined strategy set, and Miller [41] has argued that genetic algorithms can evolve strategies in a more open-ended fashion than dynamic programming. Evolutionary game theory and dynamic programming should prove useful adjuncts to differential game theory as ways of analyzing simple P-E conflicts, but the next section suggests that simulated evolution may be required to deal with complex cases.

3.3 Reasons to simulate pursuit-evasion games

Games are characterized by various dimensions of complexity: (1) the number of players, ranging from one-player cases (covered by control theory) to classic two-

player cases to more difficult multi-player cases; (2) the number of moves, ranging from "static" games of one discrete move per player (e.g. Rock, Paper, Scissors) to games with multiple discrete moves per player (e.g. chess), to differential games with continuous moves (e.g. air combat); (3) the payoff structure, with zero-sum games usually simpler than non-zero-sum games; (4) the information structure, with games of complete information much simpler than games of incomplete information. Moreover, in differential games with continuous dynamics, the complexity and noisiness of the dynamics has a major influence on the tractability of the game. Anything that complicates the differential state equations complicates the game analysis. Finally, formal application of game theory requires the complete specification of a strategy space. Such a complete specification may not be possible if the strategies are emergent properties of human heuristics, animal brains, or evolved robot control systems. These problems suggest that differential P-E games are difficult to analyze even under the best circumstances, and that the introduction of realistic complexity renders most of them formally intractable.

To avoid these complexities, differential game theory usually assumes that the P-E game is one of perfect information between two players with fixed and predetermined roles (one "pursuer" and one "evader"), deterministic dynamics and constant speeds, and a zero-sum payoff structure. Mathematically adept researchers can relax one or two of these assumptions at a time to derive results for special and simplified cases, but relaxing all the assumptions at once makes the game hopelessly complex. Some recent work attempts to analyze more difficult asymmetric and symmetric games with noise-corrupted environments [54], uncertain environments [13], or uncertain dynamics [17]. Yet even with bounded uncertainties in dynamics, the classical game-theoretic concepts of optimality, value, and saddle point may be irrelevant [17]. P-E games that cannot be reduced to differential state-space equations cannot be analyzed using the traditional methods of differential game theory. For example, without a linear and deterministic mapping from control to state variables (e.g. from a player's sensors to its effectors), it is impossible to construct tractable differential equations that relate player strategies directly to changes in the game's state-space.

Another important assumption, very rarely mentioned in game theory, is that strategies can be implemented instantaneously, without time-lags, computational costs, or speed-accuracy trade-offs. That is, decision dynamics are assumed to be much faster than behavioral dynamics. For real animals and robots, this assumption is unrealistic. Indeed, the basic assumption in game theory that unpredictability is only useful given incomplete information assumes that decision-making happens so much faster than action, that the dynamics of information-

processing are irrelevant to the dynamics of action. But if we view both cognition and action as dynamical processes operating on similar time scales [40], then the utility of unpredictability becomes more apparent. The terms “perfect” and “imperfect” information conflate the objective information structure of the game (e.g. the state-information available in the world) with the sensory and information-processing capacities of the players. If the latter are limited, then confusion, uncertainty, deception, and protean behavior may prove relevant even if the objective information structure of the game is “perfect”.

In recognition of these problems, some game theorists have recently shifted to numerical and simulation methods to derive near-optimal strategies for more complex P-E games (e.g. [29, 47]). For example, [47] used artificial intelligence (AI) methods to simulate players in an air combat maneuvering scenario. But such methods for controlling autonomous agents tend to become hopelessly slow as the dynamics of agents and environments become more complex and noisy. We need simulation methods that yield reactive, robust, dynamic P-E strategies, rather than slow, brittle, hand-designed AI systems.

Differential game theory provides a framework for describing the important features of P-E contests, and a set of normative results concerning optimal strategies in simple cases. However, it cannot generally provide optimal strategies for realistically complex P-E problems, nor can it show how strategies can be implemented in a real control system subject to limited sensory capacities, sensory and motor noise, component failure, and constraints on processing speed and accuracy. Evolutionary simulation methods can fulfill these goals and can complement game-theoretic approaches, because adaptive P-E strategies can be evolved in contest scenarios that defy formal analysis. Others have recognized this, and so we next review previous simulation work related to P-E issues.

4 Review of Related Simulation Work

Themes of pursuit and evasion are implicit in much of the recent work in artificial life and simulation of adaptive behavior. Classic problems of obstacle avoidance and of foraging and navigation can be viewed as degenerate special cases of evasion and pursuit, respectively, with the “opponents” consisting of inanimate, non-moving obstacles, food items, or other goal objects. Much of the work on simulation of collective behavior involves issues of dynamical interaction with other agents that may be similar to the those arising in P-E contests. For example, the cooperative behaviors of following, flocking, and aggregation are similar to pursuit behaviors; others such as disperson and collision-avoidance are more similar to evasion behaviors. (But note that selection for cooperation rarely favors deception or protean behavior.)

Previous simulation work has examined the origins and effects of P-E tactics with neither player evolving

or with one player evolving; these will be reviewed in order (we are not aware of any prior work with both players co-evolving.) Given very simple, fixed rules for individual movement, Schmieler [48] examined the different P-E dynamics that result when a number of simulated males and females are attracted or repulsed by one another with varying strengths, and with varying degrees of mutual knowledge about one another’s movements. Some artificial life simulations have successfully used pre-programmed predators that impose selection for simple evasion behaviors (e.g. [1, 51]).

Grefenstette’s [19, 20] SAMUEL system, resembling a classifier system, evolved robust rule-based strategies for simulated agents with noisy, coarse-grained sensors and effectors, including both effective evasion rules given one or two pre-programmed pursuers (in the ‘predator-prey’ problem), and effective pursuit rules given a randomly moving evader (in the ‘cat-and-mouse’ problem). However, SAMUEL uses high-level sensory input (e.g. direct heading, bearing, speed, and range information), symbolic condition-action rules operating in discrete time-slices (e.g. 2 to 20 decisions per contest), and fairly domain-specific genetic operators (such as Lamarckian rule deletion, generalization, and specialization).

Koza’s [30, 31] genetic programming work includes the widest array of P-E simulations. His Pac-Man scenario [31] required both evasion (of pre-programmed “monsters”) and pursuit (of sluggishly moving “fruit”); control systems evolved through genetic programming that were capable of prioritizing these activities appropriately. Some evolved Pac-Man controllers were skillful enough to eat the monsters after eating a special “pill” that made the monsters vulnerable, so to some extent the roles of pursuer and evader could be switched in this scenario. Koza [31] also investigated the evolution of P-E strategies in Isaacs’ [27] “squad car game”, where a police squad car pursues a slower pedestrian evader on a discrete grid. Most relevantly, Koza [30] used genetic programming to evolve LISP S-expression controllers for both players in a differential P-E game. His game had pre-determined and fixed roles for pursuer and evader, constant speeds for both agents, and perfect information. Agents were randomly placed in a planar world, controlled their directions (velocity vectors) based on simple inputs concerning the current heading of the opponent, and received fitness payoffs for effective pursuit or evasion. This game has a single optimal pursuit strategy (move directly towards the evader) and a single optimal evasion strategy (move directly away from the pursuer). Given an optimal evader as the “environment”, genetic programming was successful in evolving a near-optimal pursuer within a few generations (e.g. 51 generations of 500 individuals each); likewise, near-optimal evaders evolved given fixed optimal pursuers as the environment. However, Koza’s P-E game was very simple: it required mapping a sin-

gle input (current angle of opponent) onto a single output (direction to move in), given perfect information and trivial movement dynamics, and it had a known optimal solution from differential game theory. Co-evolution between pursuer and evader did not occur. Nevertheless, Koza's work represents an important fore-runner and inspiration for our research.

Aside from explicit P-E research, some simulated evolution has demonstrated the adaptiveness of mixed strategies and protean behavior. Koza [31] used genetic programming to evolve random-number generation programs under "entropy-driven evolution"; this direct selection for randomness is analogous to the indirect selection for unpredictable evasion that occurs in P-E contests. Other simulation work has shown the utility of co-evolution in evolving strategies for game-like interactions. Work by Hillis [23] on the co-evolution of sorting strategies and test sets can be viewed as an abstract version of a one-play P-E contest, in which the sorting strategies 'pursue' optimal sorts while the test sets 'evade' the strengths of particular sorting algorithms. Co-evolution has also been used successfully in Holland's [24] ECHO system, Koza's [30, 31] genetic programming research, and J. H. Miller's [43] work on the iterated prisoner's dilemma. This previous work on evolving pursuit and evasion strategies, together with Koza's demonstration of entropy-driven evolution through selection for randomness, and demonstrations of co-evolution by Hillis and others, gave us hope that a co-evolutionary P-E scenario could lead to the evolution of protean behaviors.

5 Experimental Methods and Results

In previous work on evolutionary robotics [10], Cliff *et al.* have used simulated evolution through natural selection to design sensory and control systems capable of guiding simulated robots to perform simple homing and guidance tasks. Here, as a natural extension to this work, we increase the number of agents in the world from one to two, and set up a fitness function that rewards hostile pursuit and effective evasion.

To limit computational costs, the full 3-D simulation system employed in [10] was simplified to yield a 2-D 'flatland' simulation; nevertheless, both time and space were still modeled as continuous values. Visual sensing was modeled using computer graphics techniques, with visual sensors (e.g. number, placement, and angular sensitivities of 'eyes') evolving under genetic control. All agents had the same motor system (2 wheels) and kinematics, which were modeled as differential-drive steering systems. The faster an agent moved, the larger its minimum turning radius became. All agents had the same visual appearance with distinguishable head and tail ends, so an agent could in principle detect whether its opponent was oriented towards or away from it.

Since P-E games unfold in continuous space and time

as an interplay between each agent's continuous dynamical trajectory, rather than as a series of discrete, alternating moves, we preferred to use agent control systems based on dynamical, recurrent neural networks, rather than controllers with discrete condition-action rules, such as classifier systems (e.g. [52]), LISP S-expressions [30], or deterministic finite-state automata [12]. Our agents' control systems are genetically-specified continuous-time noisy recurrent neural networks with heterogeneous time-constants, modeled using numerical approximation techniques with the same very short time-slice interval (Δt) as was used for resolving sensor and motor responses. The networks can, to a good approximation, generate continuous output based on continuous input, and have intrinsic dynamics that can be used to guide complex adaptive behavior.

In each P-E contest, a pair of individuals are placed at random on the planar surface, each moves as directed by its evolved control system based on its visual input, and each amasses fitness points for effective pursuit and/or effective evasion, depending on the fitness function. Because contest outcomes are noisy, each individual takes part in a number of contests (typically 8 or 16) to determine overall fitness. Elitist rank-based reproduction, with crossover and mutation, are used to form the next generation.

We have implemented two types of simulation. In inter-population asymmetric P-E co-evolution, a 'pursuer' population is selected for pursuit ability (like a predator species), and co-evolves against a second, reproductively separate 'evader' population, which is selected for evasion ability (like a prey species). In intra-population symmetric P-E competition, individuals in a single population compete against each other and are selected for both pursuit and evasion abilities using a single zero-sum evaluation function. The following discussion considers this latter case, which we expected to be easier, but which proved quite tricky.

We found that many of our intra-population zero-sum evaluation functions, which were designed to reward both pursuit and evasion capacities in the same individual, resulted in pseudo-cooperative solutions, such as: (1) both competitors turning away from each other and running off at full speed; (2) both competitors turning to face each other and then shutting down all motor activities, resulting in an indefinite face-off; or (3) both competitors turning to face each other and then accelerating to a high-speed collision. In most evaluation functions we have examined, the first two outcomes appear to be strong attractors, and the second two are perhaps evolutionarily stable strategies. These outcomes, and the difficulty of constructing fitness functions for intra-population co-evolution, are discussed further in [11].

In co-evolution based on zero-sum contests within populations, there are further problems in measuring real

performance increases in the population, because average population fitness will necessarily hover around zero. Figure 1 (solid line) shows the fitness score of the best individual (highest average score over 16 trials) in the population at the end of each of 300 generations. These results do not appear promising: despite selecting for *higher* scores (i.e. maximizing), fitness rapidly falls towards zero and stays there for the duration of the experiment. However, when the best individual from each generation is tested against the best individual in the original random population of generation 0 (Fig. 1, broken line), it is clear that improvement has occurred. This improvement rapidly approaches an asymptote (within 10 generations), and is largely due to genetic convergence (this experiment was performed with a high selection pressure and a relatively low mutation rate). Despite the apparent asymptote, scores vs. the best individual from generation 300 (Fig. 1, dotted line) indicate that beneficial mutations continue to occur: e.g. at generations 70 and 202; after each such mutation, performance reaches a significantly higher plateau; see [11] for further details.

By analogy, ancient predators may have caught ancient herbivores only half the time, and modern cheetahs may catch modern gazelles only half the time, but modern cheetahs would catch ancient herbivores very efficiently. To register progress in zero-sum co-evolutionary situations, we need to explicitly test individuals against their ancestors (or the ancestors of their competitors).

6 Conclusions

Pursuit and evasion behaviors are common because conflicts of interest over approach and avoidance are common, and they are difficult because dynamic, stochastic, continuous-space, continuous-time, zero-sum games are difficult. This paper has argued that the exploration of P-E contests is the next logical step in the simulated evolution of adaptive behavior. Such contests introduce many complexities, such as co-evolution, protean behavior, dynamical behavior, and collective movement patterns. We conclude by examining the engineering and scientific benefits of pursuing rather than evading these complexities.

Many traditional robot control tasks are degenerate special cases of P-E problems: collision-avoidance is evasion of non-moving obstacles, goal-directed navigation and homing behavior are pursuit of a non-moving target region, and grasping can be pursuit of a non-moving target object. Clearly, the avoidance, pursuit, or manipulation of active mobile agents radically increases the difficulty of such tasks, and their robust solution may require co-evolutionary design methods, where robot control systems evolve against pursuer or evader agents. Even where a robot's operating environment is expected to contain only static or passively moving objects, co-

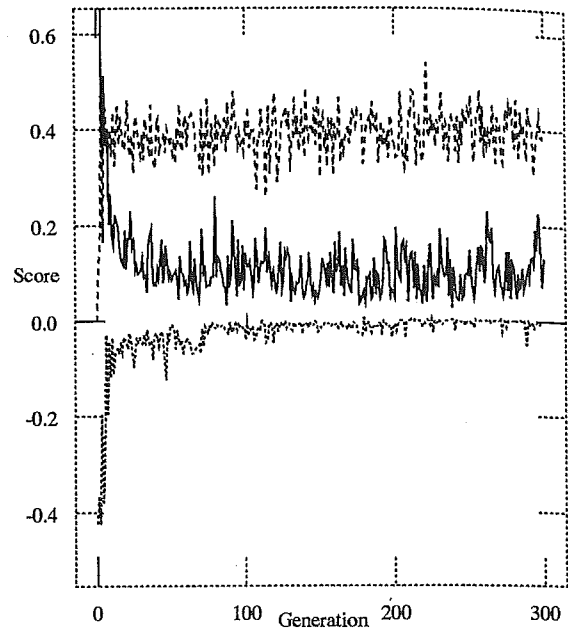


Figure 1: Scores from zero-sum P-E intra-population co-evolution. Solid line: score of best individual in generation when tested against a random selection of other individuals in that generation (peak value is 1.16 at generation 1). Broken line: score of best individual in generation when tested against best individual from generation 0. Dotted line: score of best individual in generation when tested against best individual in generation 300. All points are averages of 16 trials.

evolution of grasping and movement tactics against animate opponents might increase the robustness, speed, smoothness, and generality of control systems, because the opponents would evolve to exploit any instabilities or weaknesses in the control system. For example, a legged robot that is harried by a hostile predator that keeps trying to trip it up will probably evolve more robust walking abilities than one that merely clambers over passive obstacles. Any real robot that operates in public spaces that contain potentially hostile (or merely clumsy and curious) agents, such as children, dogs, or Luddites, must have escape and evasion abilities at least as effective as those of the average house pet. The co-evolution of evasion tactics through interaction with simulated animal, human or vehicle pursuers may help to solve this potentially catastrophic problem. Pursuit and evasion capacities have other obvious applications in computer animation, video games, and virtual environments, aside from less savory military uses.

The scientific benefits of a better understanding of pursuit and evasion would extend to game theory, animal biology, evolutionary psychology, and neuroethology. Implications also arise for our understanding of the general relationship between agents and environments. For those seeking a general theory of environmental complexity (e.g. [53, 50]), the addition of animate agents capable of unpredictable pursuit and evasion in the environ-

ment represents a significant conceptual challenge. For example, an environment that contains creatures with continuous-time dynamical recurrent networks as their control systems would be difficult to model as an environmental finite state machine, as proposed in [53]. As in sexual selection [37, 42] and other forms of "psychological selection" [38, 40], P-E contests break down the distinction between environment complexity and agent complexity, because agents become the most important selective forces in each other's environments.

Acknowledgements:

Miller's research was supported by NSF-NATO Post-Doc Fellowship RCD-9255323 and NSF Research Grant INT-9203229. Thanks to Inman Harvey, Alasdair Houston, Phil Husbands, Peter Todd, and an anonymous reviewer for useful comments and criticisms.

References

- [1] D. Ackley & M. Littman. Interactions between learning and evolution. In C. G. Langton *et al.*, eds, *Artificial Life II*, pp. 487-509. Addison-Wesley, 1992.
- [2] B. Alstermark, *et al.* Effect of different spinal cord lesions on visually guided switching of target-reaching in cats. *Neurosci. Res.*, 5(1):63-67, 1987.
- [3] M. A. Arbib & A. Cobas. Schemas for prey-catching in frog and toad. In [35, pp.142-151], 1991.
- [4] L. D. Berkovitz. Two person zero sum differential games: An overview. In J. D. Grote, ed, *The theory and application of differential games*, pp. 12-22. D. Reidel, 1975.
- [5] R. Byrne & A. Whiten, eds. *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford U. Press, 1988.
- [6] J. Camhi. *Neuroethology: Nerve cells and the natural behavior of animals*. Sinauer, 1984.
- [7] J. Camhi. Invertebrate neuroethology: Escape behavior in the cockroach and distributed neural processing. *Experientia (Basel)*, 44(5):401-408, 1988.
- [8] M. R. A. Chance & W. M. S. Russell. Protean displays: a form of allaesthetic behavior. *Proc. Zoological Soc. of London*, 132:65-70, 1959.
- [9] C. W. Clark & R. C. Ydenberg. The risk of parenthood I. General theory and applications. *Evolutionary Ecology*, 4(1):21-34, 1990.
- [10] D. Cliff, I. Harvey, & P. Husbands. Explorations in evolutionary robotics. *Adaptive Behavior*, 2(1):73-110, 1993.
- [11] D. Cliff & G. F. Miller. The co-evolution of pursuit-evasion tactics in simulated robots. Cognitive Science Research Paper, University of Sussex, 1994. In Preparation.
- [12] R. J. Collins & D. R. Jefferson. Representations for artificial organisms. In [35, pp.382-290], 1991.
- [13] M. Corliss, G. Leitmann, & J. M. Skowronski. Adaptive control for avoidance or evasion in an uncertain environment. In [55, pp.1-11], 1987.
- [14] P. Driver & N. Humphries. *Protean behavior: The biology of unpredictability*. Oxford U. Press, 1988.
- [15] A. Ermolov, B. Kryakovskii, & E. Maslov. Differential game with mixed strategies. *Avtomatika i Telemekhanika*, 47(10):1336-1349, 1986.
- [16] I. Forte & J. Shinar. Improved guidance law design based on the mixed-strategy concept. *J. of Guidance, Control, and Dynamics*, 12(5):739-745, 1989.
- [17] E. Galperin & J. Skowronski. Pursuit-evasion differential games with uncertainties in dynamics. In [55, pp.13-35], 1987.
- [18] W. Getz & M. Pachter. Two target pursuit-evasion differential games in the plane. *J. Optim. Theory Applic.*, 34:383-404, 1981.
- [19] J. J. Grefenstette. The evolution of strategies for multiagent environments. *Adaptive Behavior*, 1(1):65-89, 1992.
- [20] J. J. Grefenstette, C. L. Ramsey, & A. C. Schultz. Learning sequential decision rules using simulation models and competition. *Machine Learning*, 5(4):355-381, 1990.
- [21] W. Grimm & K. Well. Modelling air combat as differential game: Recent approaches and future requirements. In *Proc. 4th Int. Symp. Differential Games and Applications*, pp. 1-13. Springer, 1991.
- [22] W. D. Hamilton, R. Axelrod, , and R. Tanese. Sexual reproduction as an adaptation to resist parasites: A review. *Proc. Nat'l Acad. Sciences (USA)*, 87(9):3566-3573, 1990.
- [23] W. Hillis. Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D*, 42:228-234, 1990.
- [24] J. Holland. Echo: Explorations of evolution in a miniature world. In C. G. Langton *et al.*, eds, *Artificial Life II*. Addison-Wesley, 1992.
- [25] A. Houston & J. McNamara. Singing to attract a mate: A stochastic dynamic game. *J. Theoretical Biology*, 129:57-68, 1987.

- [26] A. Houston & J. McNamara. A framework for the functional analysis of behavior. *Behavioral and Brain Sciences*, 11(1):117-164, 1988.
- [27] R. Isaacs. *Differential games*. John Wiley, 1965.
- [28] R. Isaacs. The past and some bits of the future. In J. D. Grote, ed, *The theory and application of differential games*, pp. 1-11. D. Reidel, 1975.
- [29] B. Jarmark. On closed-loop controls in pursuit-evasion. In [55, pp.157-166], 1987.
- [30] J. Koza. Evolution and co-evolution of computer programs to control independently-acting agents. In [35, pp.366-375], 1991.
- [31] J. Koza. *Genetic Programming: On the programming of computers by means of natural selection*. MIT Press, 1992.
- [32] P. Maes, ed. *Designing autonomous agents: theory and practice from biology to engineering and back*. MIT/Elsevier, 1990.
- [33] J. Maynard Smith. *Evolution and the theory of games*. Cambridge U. Press, 1982.
- [34] J.-A. Meyer, H. L. Roitblat, & S. W. Wilson, eds. *Proc. Second Int. Conf. Simulation of Adaptive Behavior (SAB92)*, MIT Press/Bradford Books, 1993.
- [35] J.-A. Meyer & S. W. Wilson, eds. *Proc. First Int. Conf. Simulation of Adaptive Behavior (SAB90)*, MIT Press/Bradford Books, 1991.
- [36] A. Merz. To pursue or to evade that is the question. *J. Guidance, Control, and Dynamics*, 8:161-166, 1985.
- [37] G. F. Miller. *Evolution of the human brain through runaway sexual selection: The mind as a protean courtship device*. PhD thesis, Stanford University Psychology Department, 1993.
- [38] G. F. Miller. Psychological selection in primates: The evolution of adaptive unpredictability in competition and courtship. In A. Whiten & R. W. Byrne, eds, *Machiavellian Intelligence II*. In Press.
- [39] G. F. Miller & D. Cliff. Protean behavior in dynamic games: Arguments for the co-evolution of pursuit-evasion tactics. Cognitive Science Research Paper 311, University of Sussex, 1994.
- [40] G. F. Miller & J. J. Freyd. Dynamic mental representations of animate motion: The interplay among evolutionary, cognitive, and behavioral dynamics. Cognitive Science Research Paper 290, University of Sussex, 1993. Submitted for journal publication.
- [41] G. F. Miller & P. M. Todd. Let evolution take care of its own. *Behavioral and Brain Sciences*, 14(1):101-102, 1991.
- [42] G. F. Miller & P. M. Todd. Evolutionary wanderlust: Sexual selection with directional mate preferences. In [34, pp.21-30], 1993.
- [43] J. H. Miller. The co-evolution of automata in the repeated prisoner's dilemma. Report 89-003, Santa Fe Institute, 1989.
- [44] J. von Neumann & O. Morgenstern. *Theory of games and economic behavior*. Princeton U. Press, 1944.
- [45] M. Ridley. *The red queen: Sex and the evolution of human nature*. Viking, 1993.
- [46] E. Y. Rodin. Pursuit-evasion bibliography, version 2. *Computers and Mathematics with Applications*, 18(1-3):245-320, 1989.
- [47] E. Y. Rodin, et al. Artificial intelligence in air combat games. In [55, pp.261-274], 1987.
- [48] R. W. Schmieder. A knowledge-tracking algorithm for generating collective behavior in individual-based populations. In *Pre-proceedings Second Euro. Conf. Art. Life*, pp. 980-989, 1993.
- [49] J. Shinar, I. Forte, & B. Kantor. Mixed strategy guidance (MSG) - a new high performance missile guidance law. In *Proc. 1992 Amer. Control Conf. (Vol. 2)*, pp. 1551-1555. American Automatic Control Council, 1992.
- [50] P. M. Todd & S. W. Wilson. Environment structure and adaptive behavior from the ground up. In [34, pp.11-20], 1993.
- [51] G. M. Werner & M. G. Dyer. Evolution of herding behavior in artificial animals. In [34, pp.393-399], 1993.
- [52] S. W. Wilson. Classifier systems and the animat problem. *Machine learning*, 2:199-228, 1987.
- [53] S. W. Wilson. The animat path to AI. In [35, pp.15-21], 1991.
- [54] Y. Yavin. A stochastic two-target pursuit-evasion differential game with three players moving in a plane. In [55, pp.141-149], 1987.
- [55] Y. Yavin & M. Pachter, eds. *Pursuit-evasion differential games*. Pergamon Press, 1987.