Extended Abstract for:

Surprising Ramifications of the Surprise Exam Paradox: Exploring Rationality

Steven Orla Kimbrough
University of Pennsylvania
565 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, PA 19104-6340 USA
kimbrough@wharton.upenn.edu

February 24, 2003

Contents

> "My sensations as we approached what I supposed might be a 'field of battle' were anything but agreeable. I had been in all the engagements in Mexico that it was possible for one person to be in; but not in command. If someone else had been colonel and I had been lieutenant-colonel I do not think I would have felt any trepidation. . . . As we approached the brow of the hill from which it was expected we would see the enemy. . . my heart kept getting higher and higher until it felt as though it was in my throat. I would have given anything to have been back in Illinois, but I had not the moral courage to halt and consider what to do; I kept right on. When we reached a point from which the valley below was in full view I halted. The place where the Confederates had been encamped was still there but the troops were gone. My heart resumed its place. It occurred to me at once that [Colonel Thomas] Harris had been as much afraid of me as I had been of him. This was a view of the question I had never taken before; but it was one I never forgot afterwards. From that event to the close of the war, I never experienced trepidation upon confronting an enemy, though I always felt more or less anxiety. I never forgot that he had as much reason to fear my forces as I had his. The lesson was valuable."
>
> —Ulysses S. Grant, *Memoirs*

Paradoxes are often useful for the insights they yield when investigated. Indeed, insights are often forthcoming from paradoxes having only failed resolutions [45]. The hoary Surprise Exam Paradox (aka: Surprise Hanging Paradox) is a case in point. After rehearsing the paradox, I shall offer a new solution, or rather analysis, and begin to explore its surprising ramifications.[1]

1 The Surprise Examination Paradox

The scene is a classroom. The teacher announces that there will be a surprise examination given at one of the following six meetings of the class. The exam will be a surprise in the sense that on the morning of the day of the exam the students will not have enough information to know that the exam will occur that day with a probability at or above l, a given threshold (level or line), say $l \geq \frac{7}{10}$. The students find this puzzling and they reason as follows. "There cannot be a surprise exam on the sixth, the last, day, since we would know that morning that the exam had not yet taken place and hence that it would have to be given that day. But if we know that the surprise exam cannot happen on the sixth day, by similar reasoning it cannot happen on the fifth day. If the morning of the fifth day arrives without our having had the exam, we would know that it has to be given then, since it can't happen on the sixth day. Therefore, the exam cannot be given on the fifth day either. By continuing this reasoning process we find that it is impossible for the teacher to give us a surprise exam during the next six meetings of the class."

The teacher ignores such reasoning, privately rolls a fair die with a determination to give the exam on the day corresponding to the number that turns up. In fact a 2 results from the toss, and the teacher gives the exam on the second day. Naturally, the students are surprised. Surprise becomes distress when they see the first question on the exam:

1. Explain the fallacy in the reasoning that led you to believe it impossible for me to give you a surprise exam as announced.

How should *we* answer the examination question? There are many possible interpretations of the paradox, and for each sufficiently clear interpretation a possible, and possibly distinct, answer. My purpose is to make a particular point, and I am content to assume, without much additional explication, the interpretation needed to make the point. Here is the point.

If the teacher had rolled a 6 on the die and given the exam on the last day, indeed the students would not have been surprised. They could, let us grant, correctly have inferred the morning of the exam that it would with certainty be given that day. Further, if the die had come up a 5, the students could have correctly inferred on the morning of the fifth day that there would be a 50% chance of having the exam that day.[2] Not enough for a surprise. If any other number turned up on the teacher's roll of the die, moreover, the students could not infer with sufficient confidence to avoid surprise that there would be an exam on the day in question. The teacher's decision regime affords her a $\frac{5}{6}$ chance of giving a surprise exam. All this is obvious.

The observation to be made is that the teacher has a goal of giving a surprise exam. Let us assume that some announcement has to be made. Given this, if the teacher is unwilling to accept *any* risk of speaking falsely (of not giving the exam or of not having it be a surprise), then the teacher cannot—let us assume for the sake of the argument—offer a surprise exam. If, instead, she is willing to take on some risk—here it is a $\frac{1}{6}$ chance of no surprise—then she can possibly, even probably, achieve her goal. The backwards induction reasoning fails at the start: the exam be

[1] New so far as I know. The literature is voluminous. A sample: [5, 10, 28, 47, 50, 52]. There is a useful bibliography at http://www.magnolia.net/~leonf/paradox/hanging.txt.

[2] Could have if they knew or assumed the teacher's procedure for picking the day.

could given on the last day, it just wouldn't be a surprise. The reasoning fails to account for the teacher's willingness to assume risk for the sake of achieving her goal. Not only can the teacher do this, but she may well be entirely rational in doing so (given her tolerance of risk in the context of the prospective rewards, positive and negative).

Had the teacher announced her randomization mechanism for picking the day of the exam and said "Probably, we will have a surprise exam," no one would have cried "Paradox!". Had the teacher announced her randomization mechanism for picking the day of the exam and said "We will have a surprise exam," it would have been apparent—and not paradoxical—that the teacher stood a chance of speaking falsely and losing a "reputation bet". That situation is not changed by the teacher keeping silence on her randomization mechanism. There *can* be a surprise exam, but not with certainty. If the teacher is willing to undertake the risk of speaking falsely with the same probability of there not being a surprise exam, she can speak truly, with the same certainty with which the surprise exam materializes.

2 Finitely Iterated Prisoner's Dilemma

There is a second question on the exam:

> 2. A 100-shot Iterated Prisoner's Dilemma (IPD) game is to be played between the teacher and an unknown, but fully competent human subject. The teacher announces that she will gain the reward from mutual cooperation at least 2 times, net. That is, if P is the penalty for mutual defection and R is the reward for mutual cooperation, the teacher is asserting that she will get at least $98 \cdot P + 2 \cdot R$ points from the 100 trials. Can this assertion be plausibly justified? Is it plausibly rational? Why or why not?

Here is the one-shot PD game:

	C	D
C	R, R	S, T
D	T, S	P, P

We assume $T > R > P > S$ and $2R > T + S$. Intuition and experimental evidence—beginning with the original experiment [20, 21] and consistently since—tend to support the teacher's claim. A standard result in classical game theory does not. Is the behavior irrational? First, let us review the game-theoretic result.

It is a theorem, proved by mathematical induction, that in the finitely iterated Prisoner's Dilemma game, there is one subgame-perfect Nash equilibrium and it has both players always defecting.[3] Because D in Prisoner's Dilemma strictly dominates C for either player, probabilities are thought not to matter. Taking now the row player's perspective we could introduce probabilities for the column player's plays as follows, with $p_i = \mathit{prob}(C_i) =$ the probability that column chooser will cooperate (play C) with i games to go (i.e., p_1 on the last game played, p_2 on the next to last, etc.).

	C	D
C	Rp_i	$S(1 - p_i)$
D	Tp_i	$P(1 - p_i)$

[3] Binmore's [3, pages 354–5], and Fundenberg and Tirole's [26, pages 110-2] are two of many textbook sources for the proof.

(Let N be the number of games to be iterated in all. We begin with $G(N)$ and end with $G(1)$.) It is obvious that no matter what value p_i has, *even if it changes during the iterations of the game,* at any given step the row player is better off choosing the strictly dominant strategy, i.e., D. The same for the column player. Thus, at $G(1)$—the situation with 1 game to go—both players have D as a dominant strategy and should defect. At $G(2)$—the situation with 2 games to go—we solve the problem by solving $G(1)$ and noting that this leaves the first step at $G(2)$ equivalent to $G(1)$. Again, mutual defection wins out. And so it goes, inducting backwards to the beginning, $G(N)$.

Is there anything awry here? No, in the sense that the proof is correct that there is the unique subgame-perfect Nash equilibrium here of constant mutual defection. Yes, in that the explanatory force of the subgame-perfect Nash equilibrium has dissipated, which also explains its empirical failure in this case as well. The backwards induction argument, which is also a dynamic programming argument, assumes that p_i is independent of the history of play, of what happened at stages $N, N-1, \ldots, i+1$. Letting o_i be the outcome of the play at stage i, the backwards induction argument or dynamic programming model assumes that

$$p_i = prob(C_i|o_N, o_{N-1}, \ldots o_{i+1}) = prob(C_i) \quad (1)$$

If this assumption is violated, then the backwards induction argument fails. In the lingo of dynamic programming, the *Markov property*—roughly, that it doesn't matter how you get to a given stage, i, only that you are there—does *not* obtain [13, 44]. What players of finitely-repeated Prisoner's Dilemma have discovered from the first experiments on [20, 21]—viz. that their decisions at stage i can affect p_js ($j > i$)—is not allowed under the assumptions of the backwards induction argument. If by cooperating now a player can induce cooperation later by the counter-player, it may well be profitable to do so. What actual players properly do is to think *ahead*; something forbidden by the assumptions needed in the backwards induction argument. The backwards induction argument (for a single subgame-perfect Nash equilibrium of all defect) will fail if the corresponding dynamic programming model has assumptions that are in fact violated. Empirically, and as the teacher is betting, such a simple dynamic programming model will prove mistaken.

Cooperating now in hopes of inducing profitable cooperation later by one's counter-player is risky. It might not work. Then again it might. When N is large, and T and R are large, surely there comes a point at which we would not want to condemn the players for irrationality—in any ordinary sense, not forced by a theory—if they risk playing C earlier for the benefits of sustained cooperation later. Surely, how a rational player would want to play would incorporate the sizes of N and all the payoffs, and a judgment, however approximate, of how responsive the counter-player will be to plays of C and D. These and our player's willingness to tolerate risk will determine what the player does. It is a mistake—the same mistake the students made with the surprise exam—to assume that under no conditions of the game will the players undertake a risk of receiving S on even a single play of the N-repeated game.

Suppose the teacher of question 2 in the exam added that "I have a subjective probability of 0.9, based on general experience and reading of the behavioral game theory literature, that in the sort of case in which we find ourselves, normal human subjects can be induced to undertake a considerable amount of cooperative play. I believe, again based on experience and literature, that I have a good idea of how I should play in order to induce cooperative play by my counter-player. Given this probability (0.9), the values of the payoffs in question, and the fact that we will play 100 iterations of the game, I am prepared to undertake the risk of failing to achieve less than the claimed number of points." Would her claim of returns to be achieved seem implausible? Would we want to count her as irrational for defecting from uniformly playing D? Surely not, in either case.

3 Ramifications: Exploring Rationality

The received view in classical game theory[4] holds that

> A player does not need to know anything about the opponent to decide that it can never be optimal to use a strongly dominated strategy. It is irrational to use a strongly dominated strategy regardless of whether the opponent is Von Neumann or a chimpanzee. [3, page 149]

At the same time, the considerations adduced above are not unknown to mainstream game theory.

> What is important here is that game theory does not pretend to tell you how to make judgments about the shortcomings of an opponent. In making such judgments, you would be better advised to consult a psychologist than a game theorist. Game theory is about what players will do when it is understood that both are rational in some sense. [3, page 50]

Is there some *other* sense of rationality that will serve us better in accounting for what players do in games? A great merit of game-theoretic rationality (and rational choice theory) is that it is rigorously defined and amenable to precise investigation. The implication in "consult a psychologist" is that the phenomena to hand are best studied from a raw empirical point of view. The suggestion I wish to raise is that it is possible to be "rational in some sense" that is rigorously defined, that is amenable to precise investigation, that has good predictive and explanatory power, that bears normative force, and that does not lead us into paradox to the extent that game-theoretic rationality does. Specifics of this suggestion follow.

A *strategy* is a complete set of instructions for a player to play a game. A *metastrategy* is a procedure for finding a strategy. I use this terminology with deliberate reference to the more general concepts of heuristics and metaheuristics (cf. [43, 46, 54]). A heuristic is a decision policy or problem solution, with the added sense that it is feasible and warranted, if not optimal. A metaheuristic is a procedure for finding (presumably good) heuristics. Heuristics may be strategies in games, solutions to constrained optimization problems, and much else. For present purposes, we can think of metastrategies as metaheuristics in the context of games.

3.1 Metastrategy Desiderata

If we were to choose a metastrategy, what are the properties we would want it to have? Here is a starter list.

1. Tractable. It should be feasible and indeed tractable to execute the metastrategy in producing useful strategies. The computational cost should be bearable.
2. Specific. The metastrategy should in general produce a small number of strategies, ideally ranked in order of estimated value.
3. Productive. A metastrategy should be good in producing strategies that are effective in returning high value to the players playing the strategies. This may include the ability to exploit errors in play by a counter-player.
4. Robust. A metastrategy should perform well and productively against counter-players with varying skills, strategies, and abilities. In particular, it should be robust against non-optimal play.

[4] I quote from Binmore's excellent text book for the sake of brevity and convenience. Its views, as quoted, are representative.

5. Broadly scoped. A metastrategy should perform well with incomplete information, uncertainty, etc., and generally when the full assumption set of classical game theory is not available.
6. Teaching. A good metastrategy will find strategies effective at encouraging counter-players to play as desired.
7. Kantian. A good metastrategy should produce strategies that play well against each other, and hence can benefit by finding ways to self-correlate.[5]

One metastrategy is to find a Nash equilibrium strategy and play it. This metastrategy often fares poorly in light of the above list of requirements or desiderata; enough said of it for present purposes. Fictitious play (cf. [40, page 442], [3, page 409]) is another, in which a player estimates her best response based on a simple model of the counter-player's historical play. Again, I want simply to note it and draw the reader's attention to the general class of metaheuristics.

We may think of metaheuristics as divided (approximately) into two main classes [23]: *population-based search* procedures (or *evolution programs* (EPs), outlined as in Figure 1) and *local search* procedures (outlined as in Figure 2). The EP or population-based search family includes genetic algorithms [30], genetic programming [36], evolution strategies [22], evolutionary programming [24, 25], artificial immune systems (AIS) [12] (see also ICARIS: International Conference on ARtificial Immune Systems), memetic algorithms (e.g., particle swarm optimization and other forms, reviewed in [33]), ant colony optimization (ACO) [14, 41], learning classifier systems (LCS), and multi-population evolution programs or M-PEPs [34, 35]. Members of the EP (population-based) family are distinguished by how they undertake step 3 in Figure 1. The familiar simple genetic algorithm, for example, applies mutation and crossover operators to modify the population.[6]

---

1. Initialization: set parameters, policies, and solution-modifying operators for the run; generate (and maintain) a population of (candidate) solutions.
2. Determine fitnesses (not necessarily deterministically) of members of the population.
3. Modify the constitution of the population based on: (a) the evaluation (step 2), (b) the various operators, and (c) the parameters and policies of the run.
4. Iterate the process by going to step 2, and continuing until a stopping condition is reached.

Figure 1: Algorithm Outline for Population-Based Metaheuristics (Evolution Programs)

The local search family includes simple hillclimbing, simulated annealing [46], neural networks, tabu search, demon algorithms [9], extremal optimization [4], dynasearch [11], scatter search [39], GRASP (Greedy Randomized Adaptive Search Procedure) [19], and the family of reinforcement learning algorithms [31, 32, 53]. These local search heuristics are distinguished by how they treat steps 3 and 4 in Figure 2.

3.2 Desiderata and the Metastrategies

In the service of brevity:

[5] I allude here to Skyrms's notion of "a Darwinian version of Kant's categorical imperative: *Act only so that if others act likewise fitness is maximized*" [51, page 62].

[6] The replicator dynamic (cf. [27, 29, 55] and many other sources) conforms to the EP family of metastrategies, except that step 3, the modification step in Figure 1, lacks clause (b). Strategies themselves are not modified, only their frequencies in the population.

1. Initialization: set parameters, policies, and solution-modifying operators for the run, $\vec{p}$.
2. Create a new case by creating a new (candidate) solution, s.
3. Apply modification operators to s, yielding a new candidate solution, s'.
4. If $f(s', \vec{p})$ improves $f(s, \vec{p})$, set $s \leftarrow s'$.

   Note: $f(x, \vec{y})$ is a problem-specific fitness or evaluation function for the candidate solution x in the presence of run parameters $\vec{y}$. It need not be deterministic.
5. Update $\vec{p}$.
6. If the case is not done, go to step 3.
7. If the run is not done, go to step 2.

Figure 2: Algorithm Outline for Local Search Metaheuristics

1. Each of these metaheuristics may be used as a metastrategy. Investigations to date have largely been limited to reinforcement learning (e.g., [17, 18, 49]) and to genetic algorithms or genetic programming (e.g., [2, 15, 42]). Results have been favorable, e.g., behavior of human subjects in games has been effectively modelled (e.g., [17]), cooperation and efficient outcomes have been achieved (e.g., [2, 18, 49]), performance competitive with humans has been obtained (e.g., [16, 15, 42]).
2. A plausible case can be made that the metaheuristics or metastrategies listed here will often perform well with regard to the above list of desiderata, and in addition that they will often be effective at finding Pareto efficient, even optimal, outcomes. Further investigation is of course amply desired.
3. In the family of metaheuristics or metastrategies there is an essential step of creating new or modified heuristics and the step is blind, in the sense of Campbell's characterization of evolution as "blind variation with selective retention" [8]. Each of these metastrategies engages in a tradeoff between exploitation and exploration; each inherently undertakes some exploration and the attendant risk.
4. "[T]he usual view in economics that preferences can be taken as primitive and that behavior is fully rational" [48] may need revision if agents are shown to employ metastrategies resembling those listed here (especially with respect to undertaking exploration and risk). Experimental investigations tending to support something close to this "usual view in economics" (e.g., [1, 6, 7, 37, 38]) may need to be revisited; metastrategies offer competing forms of explanation.

In sum, the Surprise Exam paradox reminds us that taking risks may create possibilities and prospective rewards that are otherwise unavailable. The Surprise Exam teacher's empowerment by risk taking is mirrored in the finitely repeated Prisoner's Dilemma. The family of known metaheuristics (used as metastrategies) presents a rigorous and empirically attractive augmented notion of risk-taking rationality, one we might call an *exploring rationality,* in contrast to the received *Olympian* variety. An *adaptive* agent is responsive to experience and information. An *exploring* agent will take risks to obtain experience and information. The suggestion here is that exploring agents have open to them a richer and more effective rationality.[7]

[7]Profuse thanks to David Harlan Wood for many stimulating conversations and for comments on an earlier draft. I note that David suggests *promiscuous* for the *population-based* and *incestuous* for the *local search* families of metaheuristics. Thanks to Fred Murphy for a comment pertaining to the Markov property. This work was supported in part under NSF award number 9709548.

References

[1] J. Andreoni and J. H. Miller. Rational cooperation in the finitely repeated prisoner's dilemma: Experimental evidence. *The Economic Journal*, 103:570–85, 1993.

[2] Robert Axelrod. The evolution of strategies in interated prisoner's dilemma. In Lawrence Davis, editor, *Genetic Algorithms and Simulated Annealing*, pages 32–41. Morgan Kaufman, Los Altos, CA, 1987.

[3] Ken Binmore. *Fun and Games: A Text on Game Theory*. D.H. Heath and Company, Lexington, MA, 1992.

[4] Stefan Boettcher and Allon G. Percus. Extremal optimization: An evolutionary local-search algorithm. In Hemant K. Bhargava and Nong Ye, editors, *Computational Modeling and Problem Solving in the Networked World: Interfaces in Computer Science and Operations Research*, Operations Research/Computer Science Interface Series, pages 61–77. Kluwer, Boston, MA, 2003.

[5] Luc Bovens. The backward induction argument for the finite iterated prisoner's dilemma and the surprise exam paradox. *Analysis*, 57(3):179–86, 1997.

[6] Colin Camerer and Keith Weigelt. Experimental tests of a sequential equilibrium reputation model. *Econometrica*, 56(1):1–36, January 1988.

[7] Colin F. Camerer. Progress in behavioral game theory. *The Journal of Economic Perspectives*, 11(4):167–188, Autumn, 1997.

[8] Donald T. Campbell. Evolutionary epistemology. In Paul Arthur Schilpp, editor, *The Philosophy of Karl Popper*, volume XIV, Book I of *The Library of Living Philosophers*, pages 413–463. Open Court, La Salle, Il, 1974.

[9] Bala Chandran, Bruce Golden, and Edward Wasil. A computational study of three demon algorithm variants for solving the traveling salesman problem. In Hemant K. Bhargava and Nong Ye, editors, *Computational Modeling and Problem Solving in the Networked World: Interfaces in Computer Science and Operations Research*, Operations Research/Computer Science Interface Series, pages 155–76. Kluwer, Boston, MA, 2003.

[10] Timothy Y. Chow. The surprise examination or unexpected hanging paradox. *The American Mathematical Monthly*, pages 41–51, 1998.

[11] Richard K. Congram, Chris N. Potts, and Steef L. van de Velde. An iterated dynasearch algorithm for the single-machine total weighted tardiness scheduling problem. *INFORMS Journal on Computing*, 14(1):52–67, 2002.

[12] Leandro N. de Castro and Janathan Timmis. *Artificial Immune Systems: A New Computational Intelligence Approach*. Springer-Verlag, London, UK, 2002. Supporting Web site: http://www.cs.ukc.ac.uk/aisbook/.

[13] Eric V. Denardo. *Dynamic Programming: Models and Applications*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1982.

[14] Marco Dorigo and G. Di Caro. The ant colony meta-heuristic. In D. Corne, M. Ddorigo, and F. Glover, editors, *New Ideas in Optimization*, pages 11–32. McGraw Hill, London, UK, 1999.

[15] Garett O. Dworman, Steven O. Kimbrough, and James D. Laing. Bargaining by artificial agents in two coalition games: A study in genetic programming for electronic commerce. In John R. Koza, David E. Goldberg, David B. Fogel, and Rick L. Riolo, editors, *Genetic Programming 1996: Proceedings of the First Annual Genetic Programming Conference, July 28-31, 1996, Stanford University*, pages 54–62. The MIT Press, 1996.

[16] Garett O. Dworman, Steven O. Kimbrough, and James D. Laing. On automated discovery of models using genetic programming: Bargaining in a three-agent coalitions game. *Journal of Management Information Systems*, 12(3):97–125, Winter 1995-96.

[17] Ido Erev and Alvin E. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, 88(4):848–881, 1998.

[18] Christina Fang, Steven O. Kimbrough, Stefano Pace, Annapurna Valluri, and Zhiqiang Zheng. On adaptive emergence of trust behavior in the game of stag hunt. *Journal of Group Decision and Negotiation*, 11(6):449–467, November 2002.

[19] T. A. Feo and M. G. C. Resende. Greedy randomized adaptive search procedures. *Journal of Global Optimization*, 6:109–133, 1995.

[20] Merrill M. Flood. Some experimental games. Research Memorandum RM-789, RAND Corporation, Santa Monica, CA, 1952.

[21] Merrill M. Flood. Some experimental games. *Management Science*, 5(1):5–26, 1958.

[22] D. B. Fogel, editor. *Evolutionary Computation: The Fossil Record*. IEEE Press, Piscataway, NJ, 1998.

[23] Gary B. Fogel and David W. Corne. An introduction to evolutionary computation for biologists. In Gary B. Fogel and David W. Corne, editors, *Evolutionary Computation in Bioinformatics*, pages 19–38. Morgan Kaufmann Publishers, San Francisco, CA, 2003.

[24] L. J. Fogel. Autonomous automata. *Industrial Research*, 4:14–19, 1962.

[25] L. J. Fogel, A. J. Owens, and M. J. Walsh. *Artificial Intelligence through Simulated Evolution*. John Wiley & Sons, New York, NY, 1966.

[26] Drew Fudenberg and Jean Tirole. *Game Theory*. The MIT Press, Cambridge, MA, 1991.

[27] Herbert Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, Princeton, NJ, 2000.

[28] Ned Hall. How to set a surprise exam. *Mind*, 108:647–703, 1999.

[29] J. Hofbauer and K. Sigmund. *Evolutinary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK, 1998.

[30] John H. Holland. *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, 1975.

[31] J. Hu and M. P. Wellman. Multiagent reinforcement learning: Theoretical framework and an algorithm. In *Fifteenth International Conference on Machine Learning*, pages 242–250, July 1998.

[32] Leslie Pack Kaelbling, Michael L. Littman, and Andrew W. Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285, 1996.

[33] James Kennedy and Russell C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann Publishers, San Francisco, CA, 2001.

[34] Steven O. Kimbrough, Ming Lu, David Harlan Wood, and D. J. Wu. Exploring a two-market genetic algorithm. In W. B. Langdon, E. Cantú Paz, and et al., editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO 2002)*, pages 415–21, San Francisco, CA, 2002. Morgan Kaufmann Publishers. Available at http://opim-sun.wharton.upenn.edu/˜sok/mpep/papers/ga202.pdf.

[35] Steven O. Kimbrough, Ming Lu, David Harlan Wood, and D. J. Wu. Exploring a two-population genetic algorithm. Working paper, University of Pennsylvania, January 2003. Available at: http://opim-sun.wharton.upenn.edu/˜sok/mpep/papers/two-population-gas-techrep.pdf.

[36] John R. Koza. *Genetic Programming: On the Programming of Computers by Natural Selection*. MIT Press, Cambridge, MA, 1992.

[37] David M. Kreps, P. Milgrom, J. Roberts, and Robert Wilson. Rational cooperation in the finitely repeated prisoner's dilemma. *Journal of Economic Theory*, 27:245–52, 1982.

[38] David M. Kreps and Robert Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982.

[39] Manuel Laguna, Rafael Marti, and Karen Price Hossell. *Scatter Search: Methodology and Implementation in C.* Kluwer Academic Publishers, Boston, MA, 2003.

[40] R. Duncan Luce and Howard Raiffa. *Games and Decisions.* John Wiley, New York, NY, 1957. Reprinted by Dover Books, 1989.

[41] V. Maniezzo and A. Carbonaro. Ant colony optimization: An overview. In C. Ribeiro, editor, *Essays and Surveys in Metaheuristics*, pages 21–44. Kluwer, 2001. Available at: `http://www.csr.unibo.it/~maniezzo/papers/ANTStut.pdf`.

[42] Jim R. Oliver. Artificial agents learn policies for multi-issue negotiation. *International Journal of Electronic Commerce*, 1(4):49–88, Summer 1997.

[43] I. H. Osman and J. P. Kelly, editors. *Meta-Heuristics: Theory and Application.* Kluwer, Boston, MA, 1996.

[44] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* John Wiley & Sons, Inc., New York, NY, 1994.

[45] Willard V. Quine. The ways of paradox. In *The Ways of Paradox, and Other Essays*, pages 3–20. Random House, New York, NY, 1966. Originally published as "Paradox" in *Scientific American* vol. 206, 1962.

[46] Colin R. Reeves. *Modern Heuristic Techniques for Combinatorial Problems.* John Wiley & Sons, New York, NY, 1993.

[47] Nicholas Rescher. *Paradoxes: Their Roots, Range, and Resolution.* Open Court, La Salle, IL, 2001.

[48] Arthur J. Robson. The biological basis of economic behavior. *Journal of Economic Literature*, XXXIX:11–33, March 2001.

[49] T. Sandholm and R. Crites. Multiagent reinforcement learning in iterated prisoner's dilemma. *Biosystems*, 37:147–166, 1995. Special Issue on the Prisoner's Dilemma.

[50] Frederic Schick. Surprise, self-knowledge, and commonality. PDF on the World Wide Web, Accessed 2003-02-22.
URL: http://www.lucs.lu.se/spinning/categories/decision/Schick/index.html.

[51] Brian Skyrms. *Evolution of the Social Contract.* Cambridge University Press, Cambridge, UK, 1996.

[52] Elliott Sober. To give a surprise exam, use game theory. *Synthese*, 115:355–373, 1998.

[53] Richar S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction.* The MIT Press, Cambridge, MA, 1998.

[54] S. Voß. Metaheuristics: The state of the art. In A. Nareyek, editor, *Local Search for Planning and Scheduling*, volume 2148 of *Lecture Notes in Computer Science*, pages 1–23. Springer, Heidelberg, Germany, 2001.

[55] Jörgen W. Weibull. *Evolutionary Game Theory.* The MIT Press, Cambridge, MA, 1995.

`$Id: surprise-exam-abstract.tex,v 1.13 2003/02/25 03:35:48 sok Exp $`